Variation and Gradience in Phonetics and Phonology

FRANK KÜGLER · CAROLINE FÉRY
RUBEN VAN DE VIJVER · Editors

PHONOLOGY
& PHONETICS

MOUTON
DE GRUYTER

Variation and Gradience in Phonetics and Phonology

Phonology and Phonetics

14

Editor

Aditi Lahiri

Mouton de Gruyter
Berlin · New York

Variation and Gradience in Phonetics and Phonology

edited by

Frank Kügler
Caroline Féry
Ruben van de Vijver

Mouton de Gruyter
Berlin · New York

Mouton de Gruyter (formerly Mouton, The Hague)
is a Division of Walter de Gruyter GmbH & Co. KG, Berlin.

Library of Congress Cataloging-in-Publication Data

Variation and gradience in phonetics and phonology / edited by Frank Kügler, Caroline Féry, Ruben van de Vijver.
p. cm. – (Phonology & phonetics ; 14)
Includes bibliographical references and index.
ISBN 978-3-11-021931-9 (hardcover : alk. paper)
1. Prosodic analysis (Linguistics) 2. Grammar, Comparative and general – Phonology. 3. Language and languages – Variation. 4. Gradience (Linguistics) I. Kügler, Frank. II. Féry, Caroline. III. Vijver, Ruben Florentius Hendricus Eduardus van de.
P224.V37 2009
414'.6–dc22
2009024966

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

ISBN 978-3-11-021931-9

ISSN 1861-4191

Cover design: Christopher Schneider, Laufen.
Printed in Germany.

Contents

Introduction to Variation and Gradience in Phonetics and Phonology
Frank Kügler, Caroline Féry & Ruben van de Vijver 1

Accepting unlawful variation and unnatural classes
Jeff Mielke 17

Phonetic variation and gestural specification: Production of Russian consonants
Alexei Kochetov 43

Variation in the perception of an L2 contrast: A combined phonetic and phonological account
Silke Hamann 71

Prosodic conditioning, vowel dynamics and sound change
Ewa Jacewicz, Joseph Salmons & Robert Allen Fox 99

Variable quality of the Czech lateral liquid: A perception experiment with young Czech listeners
Šárka Šimáčková 125

Patterns of lenition in Brazilian Portuguese
Thaïs Cristófaro-Silva, Daniela Oliveira Guimarães 141

Silent onsets? An optimality-theoretic approach to French *h aspiré* words
Christoph Gabriel & Trudel Meisenburg 163

Gradient dorsal nasal in Northern German
Caroline Féry, Constance Hohmann & Katharina Stähle 185

Lexical exceptionality in Florentine Italian *troncamento*
Judith Meinschaefer 215

On the distribution of dorsals in complex and simple onsets in child German, Dutch and English
Ruben van de Vijver, Barbara Höhle & Susan Ott 247

Phonological knowledge in compensation for native and non-native assimilation
Isabelle Darcy, Franck Ramus, Anne Christophe, Katherine Kinzler & Emmanuel Dupoux 265

The erosion of a variable process. The case of n-deletion in Ripuarian and Limburg dialects of Dutch
Frans Hinskens 311

Minimal morpheme expression in Dutch dialectology
Marc van Oostendorp 351

Regional variation in intonation: Conversational instances of the "hat pattern" in Cologne German
Pia Bergmann 377

A model for the quantification of pitch accent realisation
Frank Kügler 405

List of authors 425

Index 428

Introduction to Variation and Gradience in Phonetics and Phonology

Frank Kügler, Caroline Féry & Ruben van de Vijver

1. Variation and gradience in phonetics and phonology

This book brings together researchers from sociolinguistics, phonetics and phonology to provide an overview of current issues in variation and gradience in phonetics and phonology. In doing so, it emphasizes the growing interest in gradience and variation in theoretical phonology. Variation is inherent to language, and one of the aims of phonological theory is to try and detect the mechanisms underlying variation.

There is variation at every level of phonological representation. Some of this variation controls the grammar and in some cases the grammar drives the variation. Examples of variation directing the grammar are found at the subsegmental level. This concerns acoustic, articulatory and perceptual cues for phonological features (see the contributions of Mielke, Kochetov, and Hamann) and the way in which they are organized in natural classes. Examples of grammar-driven variation are found at the segmental and the suprasegmental levels. At the segmental level there are grammatical differences in the production and perception of contextual variation of segments and in the syntagmatic constraints on the combination of segments (see the contributions of Gabriel and Meisenburg, Féry et al., Meinschäfer, van de Vijver et al., and Darcy et al.). Two contributions address tonal variation at the suprasegmental level: first in the way in which tones are mapped onto grammatical functions (Bergmann) and second in the way in which grammatical functions are mapped onto tones (Kügler).

In addition, variation may also depend on factors outside of language. It may be the result of membership in a specific group in the population, of being a speaker of a particular dialect or of belonging to a certain age group (see the contributions of Šimáčková, Cristófaro-Silva and Guimarães, Hinskens, and van Oostendorp). Yet another source of variation is the difference between a language as it is spoken today as compared to sometime in the past (Jacewicz et al.).

Traditional generative approaches have been conceived to explain categorical data that reflect the competence of the speaker. Since variation has often been considered to be due to the performance of the speaker, it has long been neglected in generative linguistics. Moreover, the format of the formalism has made it hard to deal with variation. In rule-based phonology, application of rules is obligatory as soon as their structural description is fulfilled and, as a consequence, speech sounds change all at once or not at all (Chomsky and Halle 1968); variation in the process of change or in the results has long been considered a side effect of production or perception, but certainly not of grammar itself. The study of variance has, therefore, been relegated to phonetics, sociolinguistics and psycholinguistics.

The emergence of new ways of thinking about linguistic patterns and about grammar has, however, changed our view of the role of gradience and variation. Nowadays the pattern of variability in the output of a rule or a process is examined to provide a more accurate description of language. This new conception of phonology has been facilitated by sophisticated, comfortable speech analysis programs and by the availability of large corpora. On this basis, linguists have rethought the role of variation and gradience in phonological theory. Numerous proposals are emerging that seek to explain variable data, such as probabilistic and stochastic models (for an overview see Anttila 2008). Phonology and phonetics influence each other in these models.

Phonology has a hard time apprehending variance. The generative phonological way of thinking has been essentially categorical in all its models and developments in the second half of the 20th century, but a revolution has started with the new millennium. Phonologists cannot pretend anymore that a certain rule or candidate evaluation applies following a yes or no principle. The application of a phonological process often depends on a myriad of factors, both external and internal, and phonology has begun to become stochastic. Gradience introduces variation into phonology. If a phonetic entity can be pronounced in different ways, depending on the environment, prosodic factors or dialectal influences, this ‘gradience’ may introduce ‘variation,’ which we understand as a stable state of grammar. And indeed, new surface-oriented models of phonology, like Optimality Theory and its constraints acting on the choice of candidates, have led to a profound rethinking of the phonological issues. If different constraints can influence a surface form in several ways, the step consisting in recognizing that a surface form can sometimes fulfill constraint C1, while at other times preferring to fulfill C2, is no longer so big. Stochastic OT, with its overlapping or tied constraints, has emerged, and the results of statistical analysis of variable forms in a corpus have become expressible. Hayes, Boersma, Anttila and others have started to propose OT changes which can account for different surface forms. As an example, Boersma (1998) proposes that constraints in grammar are organized on a continuous scale, and not on an ordinal scale as originally proposed for OT by Prince and Smolensky (1993/2004) and McCarthy and Prince (1993).

While phonologists have only recently started to identify variation and gradience as issues to be addressed in phonological theory itself, phoneticians have always dealt with the varying nature of speech sounds as a function of the context of occurrence. Phonetic research has been much more concerned, and for a much longer time, with the exact factors triggering variation. Theories of speech production capture the factors underlying variability and show the systematic quantitative pattern behind it. In particular, the theory of *Articulatory Phonology* (Browman and Goldstein 1992) assumes that the underlying mechanisms of speech production are invariant, variation of speech sounds being an epiphenomenon of surface phonetic structure. For instance, gestural overlapping speech is interpreted as a continuously increasing overlap of speech gestures, the gestures being invariantly present. In this theory, the factors that lead to overlapping speech — whether language external factors such as social class or speech style, or language internal factors such as stress, syllable and/or foot structure — all contribute to the output of speech but do not affect the actual speech gesture, thus the representation of speech.

Variation also forms the core issue of phonetically oriented sociolinguistic research. Sociolinguists relate contextually conditioned variation in speech to external factors such as social background and class (Kerswill 1996), gender (Labov 1990, 2001), and regional and/or dialectal background (e.g. Chambers and Trudgill 1998). Variation and gradient data thus play a central part in phonetic and sociolinguistic theorizing. Labov (1994) uses data from different groups of speakers to elaborate a theory of language change, arguing that by observing variation between different generations of speakers, linguists can observe linguistic change. This, in turn, can inform us about the nature of language. The main issues in sociolinguistic research concern the ‘triggering events’ (Labov 2002) that can be related to the diversity of language and linguistic patterns in the languages of the world. The presence or absence of contact between languages implies two different sources of change: On the one hand, convergence of linguistic features and units is expected from language contact, while diverging patterns need to be explained. On the other hand, in the absence of contact between languages, diverging linguistic patterns occur and possible converging patterns need to be accounted for in other ways (Labov 2002).

2. Structure of the book

The book is divided into three parts. The first part deals with variation at the *subsegmental level*, the second part with variation *between the segment and the phrase* and the final part studies variation *of the phrase*. We introduce these three parts in turn.

2.1. Variation below the segment

How variation affects the theory of features and how gradience and perceptual factors induce variation are questions addressed in the chapters in the first part of this book.

The study of variation at the level of the segment and below has led a number of researchers to inquire about the smallest units of phonetics and phonology. The traditional features of Jakobson et al. (1952), Fant (1960) and Chomsky and Halle (1968), as well as more recent versions of feature inventories (Clements 1985, Sagey 1986 and others), are being progressively replaced by more detailed representational systems. Flemming (1995) has drawn attention to the role of perception in phonology and shown that features need to be defined in terms of perception as well as production. Browman and Goldstein's (1992) Gestural Linguistic Model of Articulatory Phonology has shown that articulatory gestures also play a role in the production and perception of speech sounds, and can be further categorized according to the way different classes of sounds are articulated. Even natural classes for the classification of sounds and their role in phonological alternations (Clements and Hume 1995) are being repeatedly questioned. Speakers of different languages may have slightly different representations of similar sounds due to subtle differences in the phonological behavior of these sounds and/or their articulatory properties. Dialectal and diachronic changes, once better understood and analyzed, will most certainly deeply transform the relationship between phonetics and phonology. This book thus assists in a radical change in the understanding of how phonology and phonetics influence each other.

Jeff Mielke's paper *Accepting unlawful variation and unnatural classes* tackles the problem of the categorization of sounds. He uses a survey of natural and unnatural classes, as well as ‘unlawful’ variations of segments (Mielke 2008) (unlawful from the perspective of generative feature theory), and discusses examples of unnatural variation patterns: /o/ lowering in Schaffhausen German, the Kolami plural /-(u)l/ allomorphy, and Eastern Cheremis lenition of /d/. He then proceeds to illustrate unlawful variation with Edoid nasalization and with a dissimilation pattern called *jakan’e* in Russian dialects. Mielke points out that not all classes of segments involved in phonological rules or variations can be accounted for in terms of a small universal set of distinctive features, as has been proposed among others by Chomsky and Halle (1968) and Clements and Hume (1995). Mielke accounts for the data just mentioned with his Emergent Feature Theory, a phonological model in which both common and rare patterns are treated alike: as likely responses to the way languages exist and evolve.

In *Phonetic variation and gestural specification: Production of Russian consonants*, **Alexei Kochetov** tests some predictions of Browman and Goldstein’s (1992) Gestural Linguistic Model of Articulatory Phonology. This phonetic theory of speech production assumes that articulatory gestures, understood as the basic units of speech, are combined dynamically during articulation. The gestures have both a physical and a cognitive basis, they are ‘units of action and contrast,’ and they are defined as linguistically significant movements of articulators. As an example, Tongue Tip gesture is specified for Tongue Tip constriction degree (TTCD) and the Tongue Tip constriction location (TTCL). For TTCD, the descriptors are at least ‘closed’ (for stops), ‘critical’ (for fricatives) and ‘narrow’ (for approximants). Each tract variable is additionally specified for stiffness and damping ratio. With the help of a magnetic articulometer, Kochetov measured 12 Russian consonants articulated in different contexts. The results provide support for the model in general, while suggesting that certain modifications are necessary for the model to be able to generate a fuller range of attested patterns of gestural variation. The results also reveal some interesting interactions between lower-level phonetic units (gestures) and higher-level phonological categories (segments and the syllable), raising the question of how such interactions can be modeled in the framework.

Silke Hamann’s paper *Variation in the perception of an L2 contrast: A combined phonetic and phonological account* proposes an OT account that captures listener-specific variation. Perceptual confusion may arise because of a difference in the phonological categories involved in different languages: German learners of Dutch were asked to identify a threefold contrast of Dutch labiodentals, the two fricatives /f/ and /v/, and the approximant /ʋ/ (German only displays a binary contrast between /f/ and /v/). When asked to categorize different sounds, listeners apply different strategies: Advanced learners show consistently better ratings and learners at the beginner's stage consistently poorer ones. Hamann observed four different categorization patterns. She proposes that each of the four patterns can be modeled with variants of perception grammars that the learners develop according to their L2 input on the basis of their L1 grammar. In addition, learners at the beginner's stage use their L1 grammar to distinguish the sounds. The individual grammars are influenced by two factors, the auditory input of certain phonetic features and the weightings thereof. On the basis of an earlier detailed phonetic study of the duration and harmonicity median of Dutch labiodentals (Hamann and Sennema 2005), Hamann develops cue constraints to map the auditory input onto abstract phonological categories. These cue constraints are employed to formalize the listeners' choice and weighting of perceptual cues and their postulated phonological categories, which are responsible for the observed variation in L2 perception.

The overall research question of **Ewa Jacewicz**, **Joseph Salmons** and **Robert Allen Fox**'s paper *Prosodic conditioning, vowel dynamics and sound change* is the synchronic basis of vowel changes and the subject of investigation is vowel changes in two Midwestern dialects in relation to prosodic prominence. Wisconsin English is undergoing Northern Cities Shift while Ohio English is not. The authors contrast tense /e/ with lax /ɛ/ to test Labov's (1994) Diachronic Principles I and II, which state that long tense vowels have a tendency to rise in chain shifts whereas short lax vowels show a tendency to fall. In a production experiment they find that, in the Wisconsin dialect, duration, F1 and F2 of the vowels are enhanced in more emphatic environments. By contrast, the Ohio vowels /e/ and /ɛ/ are longer and more fronted relative to the Wisconsin vowels, and the longer durations of the Ohio vowels also cause a greater amount of frequency change. In the case of tense diphthongal /e/, dynamic rather than static vowel characteristics determine the nature and direction of the change. In monophthongal /ɛ/, acoustic correlates and dialectal differences have merged, though the general direction of the change confirms Diachronic Principle II. A perception experiment was conducted to answer the following question: "Would greater prosodic prominence create a clearer exemplar of a vowel or would it cause a vowel category shift?" One result is that the longer, more fronted and more diphthongal Ohio vowels were identified better by both Ohio and Wisconsin listeners. By contrast, the

Wisconsin vowels were misinterpreted as their neighbors /i/ or /ɛ/. The authors conclude that Wisconsin /e/ may be undergoing a perceptual category shift under greater enhancement.

In the paper *Variable quality of the Czech lateral liquid: A perception experiment with young Czech listeners*, **Šárka Šimáčková** investigates an ongoing sound change. The apico-alveolar lateral approximant is velarized in the speech of many young Czechs. In the speech therapy literature of Czech, such a pronunciation is frowned upon. A review of the relevant literature of the last 10 years reveals that the percentage of school children diagnosed with this problem has doubled, while other problems have remained more or less stable. This does not indicate that "the country has been hit by an epidemic," but rather that the pronunciation is becoming more variable. Šimáčková wonders whether the norms with respect to certain pronunciations are also changing. Is a velarized /l/ still perceived as sloppy? A production and perception study shows that the norms are indeed changing: Czechs prefer a light [l] prevocalically, but they accept a dark ([ɫ]) or vocalized ([o̯]) variant of the lateral post-vocalically.

In their paper *Patterns of lenition in Brazilian Portuguese*, **Thaïs Cristófaro-Silva** and **Daniela Oliveira Guimarães** discuss two connected processes that are changing the phonology of the Brazilian Portuguese of Belo Horizonte in Minas Gerais on the basis of a corpus study. The first process involves the palatalization of an alveolar sibilant [s] → [ʃ] in the context of a palatal affricate [tʃ], and the second one involves a cluster simplification: a [ʃtʃ] cluster is lenited to [ʃ]. A noteworthy aspect of these processes is that neither of them applies across-the-board, but they occur in a lexical diffusion fashion. Lenition takes place in the onset of unstressed syllables in more frequent words, and non-lenited forms, i.e. those which present an affricate, occur in the onset of stressed syllables in less frequent words. The effect of frequency on the application of the processes suggests that phonological change takes place in a gradual, word-by-word fashion, and not across-the-board.

2.2. Variation between the segment and the phrase

The next set of papers show that many phonological alternations which have been analyzed with the help of categorical rules in the past are in need of a radically different approach. These rules are in fact gradient and their applicability depends on contextual factors, which can be sociological or internal to the grammar. Traditional generative approaches are unable to account for such variations. This problem is of course not new, and was already addressed several decades ago by Labov (1969), Cedergren and Sankoff (1974) and others, who proposed the so-called variable rules, a simplified and not quantified re-write rule format à la SPE (Chomsky and Halle 1968). In these rules the contextual factors are organized in such a way that they are part of the structure of the rule itself and at the same time reflect the decreasing influence they exert on the frequency of the rule application. But these rules have not found wide acceptance in mainstream phonology. In the last ten years, Optimality Theory has been a general framework for the development of different methods for accounting for gradience and variation. Some of the papers in this part show how to integrate language-internally conditioned processes of variation into the phonological theory. The other aspect of papers in this part is concerned with external factors that govern variation.

Christoph Gabriel and **Trudel Meisenburg**'s paper *Silent onsets? An optimality-theoretic approach to French* h aspiré *words* examines a classic problem of French phonology: *h aspiré* words. These words exhibit a special behavior with respect to resyllabification phenomena, such as *enchaînement* (a coda is syllabified as the onset of the following vowel-initial syllable), *liaison* (a latent coda consonant is pronounced) and *élision* (a vowel is not pronounced in a hiatus environment). The authors show that these processes are optional in *h aspiré* words, but are nearly obligatory in other vowel-initial words. The investigation is based on a production experiment in which 12 native speakers of French produced 48 utterances. Most speakers made a distinction between *h aspiré* words and vowel-initial words, although there was a great deal of variation. It looks as if the left edge of these words is protected by a consonantal slot, which Gabriel and Meisenburg propose to be a creaky glottal stop. This consonantal slot is then realized in different ways in different environments. The second part of the paper proposes an optimality-theoretic analysis of the variation. A small number of constraints can be used which deliver the observed realizations. The paper ends with an account of variation and optionality in Optimality Theory (OT) and makes a comparison between the predictions made by Anttila's (1997, 2002) Stratified Grammars and Boersma and Hayes' (2001) stochastic Gradual Learning Algorithm (GLA). The latter approach accounts for the proportions between the variants more closely than Anttila's model.

Caroline Féry, **Constanze Hohmann** and **Katharina Stähle**'s paper *Gradient dorsal nasal in Northern German* studies the variation between two realizations of the dorsal nasal in a dialect of German, [ŋ] and [ŋk]. On the basis of data obtained in a production experiment, they show that the allomorphy is not categorical (contra much previous literature), but gradient, and that the factors determining the choice between the allomorphs are phonological in nature: prosodic boundary, accent and vowel quality. Traditional phonological models cannot account for such variation. What is needed is an extension of the stochastic Gradual Learning Algorithm (GLA) (Boersma 1998; Boersma and Hayes 2001), called the factor-driven gradient model (FDG). Instead of relying on random noise to explain temporary optionality between two allophones as in GLA, in FDG the factors bearing on the distribution of the data are part of the phonology formalism. Context-independent constraints define a baseline for the gradience, and these constraints are separated and influenced by context-sensitive constraints.

Judith Meinschaefer's paper *Lexical exceptionality in Florentine Italian troncamento* discusses the case of Italian troncamento from a new perspective. Troncamento is a syncope process that is responsible for the deletion of a word-final vowel when preceded by a sonorant. Using data from Florentine Italian, Meinschäfer argues for a systematic prosodic and morphological conditioning of this lexical exception. The data is spontaneous speech collected from an Italian speech corpus. Using this kind of data, Meinschäfer adds a new data source. She claims that the domain of application of troncamento is the phonological phrase and that its occurrence is rhythmically conditioned. The analysis of the data shows that the syncope process only applies to fully inflected word forms to the exclusion of all other word forms. An OT analysis accounts for the regularities behind this process, which has previously been claimed to be largely optional phonological variation.

In the paper *On the distribution of dorsals in complex and simple onsets in child German, Dutch and English*, **Ruben van de Vijver, Barbara Höhle** and **Susan Ott** study the acquisition of dorsals in simple and complex onsets. They discuss acquisition data from three languages, German, English and Dutch, and show that in these languages some children realize singleton dorsals as coronals (e.g. [ku] 'cow' is realized as [tu]). Dorsals are only allowed in restricted contexts: in all kinds of clusters or before unstressed vowels. The German children realize dorsals in clusters (e.g. German [klaɪt] ‘dress’ is realized target-like as [klaɪt]), but nowhere else. In Dutch and English dorsals may appear in complex onsets as well as before unstressed vowels. An example is the second dorsal in the Dutch word [kɪkər] ‘frog’, which is realized as [tɪkər]. The authors explain this pattern with two positional markedness constraints. The first one allows dorsals in complex onsets and the other one prohibits dorsals before stressed vowels. These constraints correctly account for the distribution of dorsals in the language of these children. However, they fail to make the correct typological predictions. This shows, according to the authors, that these constraints are learned.

The paper *Phonological knowledge in compensation for native and non-native assimilation* by **Isabelle Darcy, Franck Ramus, Anne Christophe, Katherine Kinzler** and **Emmanuel Dupoux** discusses the consequences of certain patterns of assimilations. Most, if not all, languages have processes of assimilation of one form or another. Nasal place assimilation is one of the most well-known types of assimilation: A nasal adopts the place of articulation of a following stop. For theories of word recognition, assimilations pose an interesting problem. All models assume that recognition takes place when some output string matches an input string. An assimilated nasal prevents such a match. To explain the fact that recognition is not hampered by an obvious mismatch between output and input, psycholinguists hypothesize a mechanism called *compensation.* There is a considerable amount of debate about whether compensation is *auditory* and *language-independent* or *phonological* and *language-specific.* The authors set up experiments in which they looked at compensation for assimilation of Place of Articulation, which is active in English but not in French, as well as regressive voicing, which is active in French but not in English. These experiments were aimed at testing the hypothesis that compensation is phonological and language-dependent. They used the mentioned processes to create stimulus material for both languages, so that it would contain an active and an inactive process in each language. Reaction times in various tasks would give clues about the well-foundedness of their hypotheses. Their reasoning was that if they are correct, the native speakers of each language would compensate for a different process. Their results confirm this hypothesis.

The erosion of a variable process: The case of n-deletion in Ripuarian and Limburg dialects of Dutch by **Frans Hinskens** looks at a phonological process from a linguistic as well as sociolinguistic point of view. Hinskens presents a study of ongoing processes of dialect leveling in a group of Ripuarian and Limburg dialects of Dutch. The feature studied is the variable deletion of final /n/ in monosyllabic words when preceded by a short vowel (as *in* 'in', *ben* 'am', *dan* 'then; than', and so on). The variability of the process is related to its complex, multiple (partly phonological, partly phonetic) conditioning. It is shown that the deletion process is blocked in a specific constellation of prosodic conditions. Language change is intimately related to variation. The paper discusses the conditions of the deletion rule and its gradual loss; n-deletion is probably in the process of changing from a dialect feature in the traditional sense to a sociolect feature. The quantitative approach that Hinskens presents provides evidence that linguistic change is underway from a geographical into a social feature.

Marc van Oostendorp takes on the problem of opacity in his paper *Minimal morpheme expression in Dutch dialectology*. Opaque processes are those which apply even if their context is not met, or which do not apply even if their context is met. He discusses two processes in various Dutch dialects. In both cases, an inflectional morpheme is not present at the surface, but it still exerts its influence. To explain these facts, van Oostendorp formulates a principle of *Phonological Recoverability*, which claims that every morpheme in the input should be recoverable in the output. This principle is formulated as an OT constraint and requires that a phonological feature in the input expression of a morphological feature must always surface. Support for this proposal comes from a detailed analysis of morphophonological phenomena in several Dutch dialects. In these dialects phonological processes are blocked if their application would make it impossible to recover the morphological input.

2.3. Variation at the level of the phrase

Phrase-level phonology is concerned with prosody and intonation and the question addressed here is to what extent postlexical tones and tunes are subject to variation. The two papers in this section deal with dialect intonation, the variation which arises between a standard language and its dialectal varieties. The crucial assumption behind the study of dialect intonation is that any dialectal variety represents its own grammar. A basic understanding of dialect grammars enables a full exploration of the variation between dialects, and thus makes it possible to capture the whole range of grammatical options a language may employ. Ladd (1996) proposes a useful tool to capture intonational differences between languages at different linguistic levels. According to Ladd these levels concern semantic, systemic, realizational and phonotactic differences; semantic and realizational aspects are of particular interest here.

Semantic differences in the intonation of languages capture the "differences in the meaning or use of phonologically identical tunes" (Ladd 1996, 199). The distinct meanings or functions conveyed by one and the same intonational tune represent a clear case of variation between language varieties. For instance, Ladd reports on high rising question intonation, which in American varieties of English is frequently used to express a kind of statement while at the same time asking for feedback from the interlocutor. For listeners of British English, Ladd reports that this particular use of the high rising question intonation sounds wheedling or insistent. In British English, this tune only conveys a meaning of request. These subtle differences between varieties of a language provide the range of possible variation of a phrase-level phonological entity.

Realizational differences in the intonation of languages capture the "differences of detail in the phonetic realisation of what may be regarded phonologically as the same tune" (Ladd 1996, 199). The realization of tones is expressed in terms of different measures. One well-studied aspect is the alignment of tones relative to the segments, syllables or tone bearing units. The alignment of tones depends on different factors such as syllable structure, proximity of other prosodic events (accent tones or prosodic boundaries), discourse structure, and information structure. However, given a situation in which these factors are constant, dialect varieties may still differ in the exact segmental anchoring of tones (Atterer and Ladd 2004; Gilles 2005). Gilles (2005) shows in a comparison of German varieties that the high tone of a falling accent is realized earlier in West German varieties than in East German varieties. Thus dialectal variation is an important factor in capturing the phonological entities of grammar.

In *Regional variation in intonation: Conversational instances of the 'hat pattern' in Cologne German*, **Pia Bergmann** examines the form and function of rise-fall contours in a corpus of spontaneous speech in the dialect of Cologne German. The proposed analysis combines standard phonological analysis of tunes in terms of the autosegmental-metrical model of intonation with methods from conversation analysis. Based on naturally occurring speech data, the functional interpretation in terms of conversational analysis captures the range of functions of the rise-fall better than those accounts found in the general phonological literature, thus providing new insights into the use of the rise-fall contour. On the formal tonological level, Bergmann identifies two distinct rise-fall contours. Both contours consist of a rising pitch accent, however the falling part of these contours differs. In one contour the falling pitch accent is downstepped, while in the other it is a simple falling accent. On the functional level the Cologne rise-fall shares the use of the Standard German 'hat pattern' (Féry 1993), namely cases of a multiple foci or topic-focus structure. As for syntax, the falling part of the contour often occurs on syntactically expanded elements, i.e. elements that are attached after a syntactic completion point. And as for information structure, many instances of the falling part are non-rhematic. The prosodic marking of non-rhematic constituents appears to be in line with the results of a study on degrees of givenness (Baumann and Grice 2006). The use of a pitch accent on constituents with different degrees of rhematicity provides further evidence for the recent claim that phonological categories do not define information structural categories in a one-to-one manner (Féry 2007).

Frank Kügler's paper *A model for the quantification of pitch accent realisation* proposes a phonetically based model for the analysis of tonal gestures to capture intonational variation in the phonetic implementation of tones between language varieties. It is assumed that the phonetic realization of tones affects two levels of implementation, the horizontal level in terms of tonal alignment and the vertical level in terms of tonal excursion. In a comparison between two German dialects, Swabian and Upper Saxon German, Kügler identifies two parameters that, taken together, allow the quantification of intonational variation with respect to tonal implementation across language varieties. The speech data used is from map task dialogues and is thus fairly uncontrolled. However, it is shown that phonologically identical contours of the two dialects in question differ significantly in terms of alignment and excursion. On the basis of this model observed differences between functionally and phonetically identical yet phonologically different contours are quantifiable.

3. Conclusion

The chapters in this book reflect the growing relevance of the study of variation and gradience for phonological theory. The availability of new theories, in which variation and gradience can be modeled, and new technical aids, which have made it clear that a more principled account of variation and gradience is necessary, has contributed to this development. We believe that these advances have strengthened phonological theory and will continue to do so.

References

Anttila, A.
1997 Deriving variation from grammar: A study of Finnish genitives. *Variation, change and phonological theory*: 35-68.

Anttila, A.
2002 Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory* 20(1): 1-42.

Anttila, A.
2008 Gradient phonotactics and the Complexity Hypothesis. *Natural Language and Linguistic Theory* 46(4): 695-729

Atterer, M. and D. R. Ladd
2004 On the phonetics and phonology of ‘segmental anchoring’ of F0: Evidence from German. *Journal of Phonetics*, 32(2): 177-197.

Baumann, S. and M. Grice
2006 The intonation of accessibility. *Journal of Pragmatics* 38(10): 1636-1657.

Boersma, P.
1998 *Functional Phonology*. Ph.D. thesis, University of Amsterdam.

Boersma, P. and B. Hayes
2001 Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32(1): 45-86.

Browman, C. and L. Goldstein
1992 Articulatory phonology: An overview. *Phonetica* 49(3-4): 155-80.

Cedergren, H. and D. Sankoff
1974 Variable rules: Performance as a statistical reflection of competence. *Language* 50(2): 333-355.

Chambers, J. K. and P. Trudgill
1998 *Dialectology*. Cambridge: Cambridge University Press.

Chomsky, N. and M. Halle
1968 *The Sound Pattern of English*. New York: Harper and Row.

Clements, G. N.
1985 The geometry of phonological features. *Phonology Yearbook* 2: 223-252.

Clements, G. N. and E. Hume
1995 The internal organization of speech sounds. In *The Handbook of Phonological Theory*, J. Goldsmith (ed.), 245-306. Oxford: Blackwell.

Fant, G.
1960 *Acoustic Theory of Speech Production*. Berlin: Mouton.

Féry, C.
1993 *German intonational patterns*. Tübingen: Niemeyer.

Féry, C.
2007 Information Structural Notions and the Fallacy of Invariant Correlates. In *Working Papers of the SFB632, Interdisciplinary Studies on Information Structure (ISIS) 6*, C. Féry, G. Fanselow and M. Krifka (eds.), 161-184. Potsdam: Universitätsverlag Potsdam.

Flemming, E.
1995 *Auditory Representations in Phonology*. Ph.D. thesis, University of California, Los Angeles.

Gilles, P.
2005 *Regionale Prosodie im Deutschen: Variabilität in der Intonation von Abschluss und Weiterweisung*. Berlin: Mouton de Gruyter.

Hamann, S., and A. Sennema
2005 Voiced labiodental fricatives or glides - all the same to Germans? In *Proceedings of the Conference on Plasticity in Speech Processing*, Valery Hazan and Paul Iverson (eds.), 164-167. London: UCL.

Jakobson, R., G. Fant and M. Halle
1952 *Preliminaries to Speech Analysis: The Distinctive Features*. Cambridge: MIT Press.

Kerswill, P.
1996 Children, Adolescents, and Language Change. *Language Variation and*
Change 8(2): 177-202.

Labov, W.
1969 The logic of nonstandard English. *Georgetown Monographs on Language and Linguistics* 22: 1-31.

Labov, W.
1990 The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2(2): 205-254.

Labov, W.
1994 *Principles of Linguistic Change: Internal Factors*. Oxford: Blackwell.

Labov, W.
2001 The anatomy of style shifting. Stylistic variation in language. In *Style*

and sociolinguistic variation, P. Eckert and J. R. Rickford (eds.), 85-108. Cambridge: Cambridge University Press.

Labov, W.
2002 Ordinary events. *A contribution to the proceedings of a conference at Pitzer College. http://www.ling.upenn.edu/wlabov/home.html.*

Ladd, D. R.
1996 *Intonational Phonology*. Cambridge: Cambridge University Press.

McCarthy, J. and A. Prince
1993 Prosodic morphology I: Constraint interaction and satisfaction. Ms. University of Massachusetts, Amherst and Rutgers University, New Brunswick [Technical Report No. 3, Rutgers University Center for Cognitive Science. To appear, Cambridge, Mass.: MIT Press].

Mielke, J.
2008 *The emergence of distinctive features*. Oxford: Oxford University Press.

Prince, A. and P. Smolensky
1993/2004 Optimality theory: Constraint interaction in generative grammar. Ms. Rutgers University, New Brunswick and University of Colorado at Boulder [Technical Report No. 2, Rutgers University Center for Cognitive Science.]

Sagey, E.
1986 *The Representation of Features and Relations in Non-Linear Phonology*. Ph.D. thesis, MIT, Cambridge.

Accepting unlawful variation and unnatural classes[1]

Jeff Mielke

1. Introduction

A widely-held view among phonologists is that there is a small universal set of distinctive features out of which segments and sound patterns are constructed. In this view, sound patterns are expected to involve classes of segments which are definable in terms of the correct set of features. Further, when related languages or dialects exhibit variations on a common sound pattern, it would make sense for these classes to be relatable in terms of features. While it is well known that some sound patterns and patterns of variation are problematic for innate features, there is no consensus as to how common they are, and therefore there is not much of a sense of how problematic they are to the versions of linguistic theory that prohibit them. This paper argues that "unnatural classes" and "unlawful variation" are both widespread. Evidence is presented in Section 2 from apparently unnatural classes found in a survey of sound patterns (Mielke, 2008) as well as cases of variation between related languages or dialects with a common sound pattern. Theories of innate features make predictions about the nature of this variation, and it is seen that the data do not bear out these predictions, that variation which is "unlawful" from the standpoint of innate features is well-attested. *Emergent Feature Theory* (Mielke, 2008) is supported in Section 3 as means of dealing with these sound patterns, and in Section 4, a simulation of phonetic generalization based on an objective model of phonetic similarity is put forth as further evidence in support of this approach.

A classic example of an "unlawful variation" pattern is found in northeastern Swiss German, where /o/ is lowered to [ɔ] before certain consonants (Keel, 1982). In and around the city of Schaffhausen, four different versions of /o/ lowering are observed, in which different classes of sounds condition lowering. Some of this variation is illustrated in (1). /o/ lowering seems to have originally occurred only before /r/, but to have been generalized differently in different communities (Keel, 1982; Janda and Joseph, 2001; Janda, 2003), as shown in Figure 1.

(1)

a. /o/ is lowered before /r/ throughout Schaffhausen.

[bɔrə]	‘to bore’	[tɔrn]	‘thorn’
[fɔrə]	‘fir tree’	[kwɔrffə]	‘thrown pp.’
[hɔrn]	‘horn’	[ʃpɔrə]	‘spur’

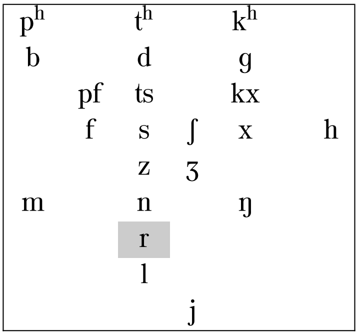
b. /o/ is lowered before coronal obstruents elsewhere in the canton, but not in the city of Schaffhausen itself.

City of Schaffhausen	Cantonal dialects	
[ʃtotsə]	[ʃtɔtsə]	‘to push down’
[losə]	[lɔsə]	‘to listen’
[ross]	[rɔss]	‘horse’
[xrottə]	[xrɔttə]	‘toad’
[got]	[gɔt]	‘god’
[ʃnodərə]	[ʃnɔdərə]	‘to stir. pp.’
[pottə]	[pɔttə]	‘offered, pp.
[ksottə]	[ksɔttə]	‘boiled, pp.’
[tnossə]	[tnɔssə]	‘enjoyed, pp.’
[kʃossə]	[kʃɔssə]	‘shot, pp.’

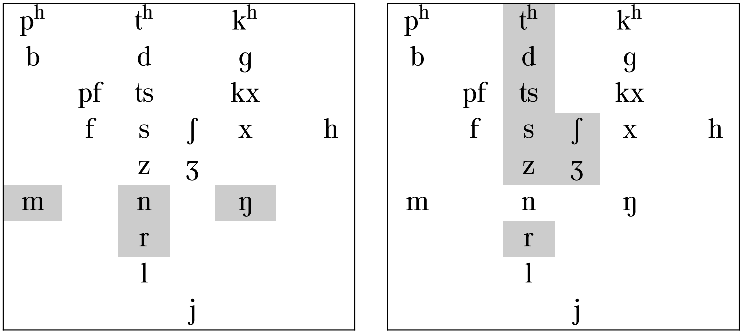
In the city of Schaffhausen, the conditioning environment for lowering includes nasals (i.e., other nonlateral sonorants) (Figure 1b), while in 17 nearby villages, the environments have been generalized to include other nonnasal, nonlateral coronal consonants (Figure 1c). In 13 other villages, the generalization includes nasals and coronal obstruents (i.e., segments which are similar to /r/ in one of two ways: sonorance and coronality) (Figure 1d). In five villages, the conditioning environment has generalized to include all obstruents except /b/ (Figure 1e).[2]

While it is probably universally expected that these classes should involve similar groups of sounds, we can imagine a scenario in which the classes are easily definable with distinctive features, and differences between classes amount to minor changes in the feature specifications, but this does not seem to be the case. Instead, the feature specifications[3] of the classes (shown in Figure 1) are quite different from each other. Furthermore, two apparent generalizations resulted in classes which are not characterizable with a conjunction of distinctive features. The overgeneralization in Figure 1d requires the union of natural classes, while Figure 1e requires subtracting one class from another, or else the union of a larger number of classes. This suggests that speakers may simply learn the environments where /o/ lowering occurs in the speech of members of their community, regardless of whether or not the set of segments involved is expressible as a conjunction of distinctive features within any particular theory, or whether the feature specification is similar to the one at the previous stage. Other factors appear to be driving these changes, and in two cases, the result is a class that can be easily described with distinctive features. It is interesting that one class (1d) is the union of two classes used for the sound pattern in nearby dialects. This supports the use of two feature bundles to describe it, but the other problematic class (1e) is not at all amenable to this approach.

In innatist theories, the classes of sounds which are eligible to participate in sound patterns are expected to be definable in terms of the innate distinctive features. For this reason, it is important to know how common cases like Schaffhausen /o/ lowering really are. If an innate set of features underlies phonological patterns, "unnatural" classes should be few and far between. The results of changes which give rise to an assortment of classes participating in basically the same sound pattern could reasonably be expected either to be limited to classes that are natural according to the feature set, and the feature specifications of the classes could be expected to be similar. If the Schaffhausen scenario is not unusual, then innate feature theory should be reevaluated. The next section investigates the occurrences of unnatural classes in a large database of sound patterns, with an eye toward cases where related languages show variations on a common sound pattern.

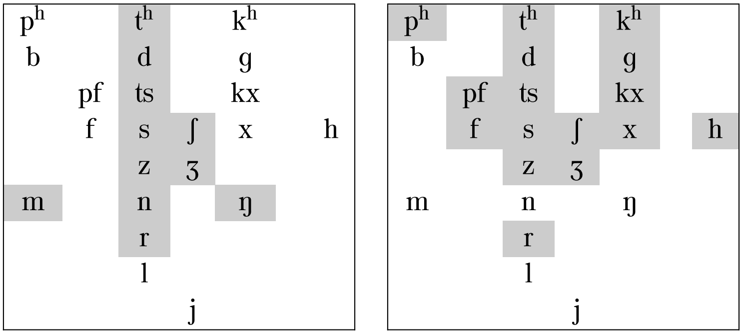


a. Apparent phonetic basis:
/o/ → [ɔ] / __r



b. Schaffhausen proper:
[+son, –voc, –lat]

c. 17 nearby villages:
[–voc, –lat, –nas, +cor]



d. 13 nearby villages:
[+son, –voc, –lat]
∨ [–voc, –lat, –nas, +cor]

e. 5 nearby villages:
[–voc, –lat, –nas]
– [+voi, +lab]

Figure 1: /o/ lowering in Schaffhausen.

2. A crosslinguistic survey of (un)natural classes

Mielke (2008) conducted a survey of classes involved in a large number of languages, in order to test predictions of innatist and emergentist feature theories. The object of the survey is the phonologically active class, defined as in (2).

(2) Phonologically active class: any group of sounds which, to the exclusion of all other sounds in a given inventory:
  a. undergo a phonological process,
  b. trigger a phonological process, or
  c. exemplify a static distributional restriction.

Phonologically active classes were collected from grammars of 561 languages (all the grammars available on the shelves in Library of Congress PA-PM at the Ohio State University and Michigan State University libraries). Looking only at the classes which undergo or trigger processes (2-a,b), there are 6077 distinct classes, some representing multiple sound patterns in a particular language.

Feature analyses of all classes were performed in three well-respected feature theories: Preliminaries (Jakobson et al., 1954), SPE (Chomsky and Halle, 1968), and Unified Feature Theory (Clements and Hume, 1995). All of the classes discussed here are of course naturally occurring, and the terms “natural” and “unnatural” are used in reference to specific feature theories, as defined in (3) and (4).

(3) Natural class: A class of sounds is natural with respect to a particular theory if the class is statable as a conjunction of features in that theory.

(4) Unnatural class: A class of sounds is unnatural with respect to a particular theory if it is not statableas a conjunction of features, but rather requires special treatment, such as union or subtraction of natural classes, or is unstatable in terms of features in the theory.

The ability of three feature systems to characterize 6077 phonologically active classes with a conjunction of distinctive features is shown in Table 1.

Table 1: The ability of three feature systems to characterize 6077 phonologically-active classes (Mielke, 2008, 2005a).

Feature System	Characterizable (natural)	
Preliminaries	3640	59.90%
SPE	4313	70.97%
Unified Feature Theory	3872	63.72%
ANY SYSTEM	4579	75.35%

Unnatural classes are not particularly rare. 1496 classes (24.65%) are unnatural in all of these feature theories. The next subsection deals with these featurally unnatural classes in more detail.

2.1. Unnatural classes

The featurally unnatural classes found in the survey can be roughly partitioned into two types: non-recurrent "crazy" classes and recurrent classes which may involve shared phonetic properties, but properties which do not happen to correspond to generally-recognized distinctive features.

An example of the former "crazy" type of class occurs in Kolami (Emeneau, 1961: 46-50). The suffix /-(u)l/ is a plural marker for a variety of nouns, and the allomorphy is phonologically conditioned. The [-l] allomorph is conditioned by /t d n̪ i iː e eː a aː/, while the [-ul] allomorph is conditioned by /p t̪ k d̪ g s v z m ŋ j/ (Table 2). Even if one allomorph is treated as basic and the other derived, there is no way to characterize the derived class in terms of traditional distinctive features, or to describe it in terms of shared phonetic properties. The most glaring reason for the unnaturalness of this class is the fact that the dental nasal patterns with the retroflex stops but not with the dental stops.

(5) Kolami plural /-(u)l/ allomorphy (Emeneau, 1961: 46-50)

a. [-l] after / t d n̪ i iː e eː a aː/

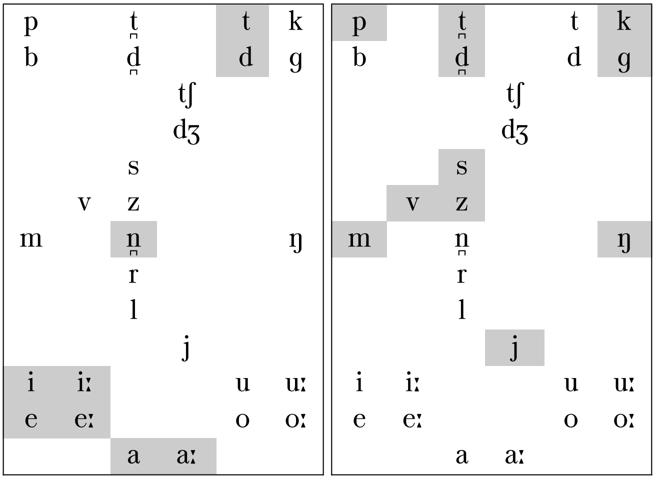
singular	plural	
d̪ut	d̪utl	'hips'
ed	edl	'bullock'
toːren̪	toːren̪l	'younger brothers'
sir	sid̪l	'female buffalo'

kaje	kayel	‘fish’
biː-am	biːl	‘rice’
kala	kalal	‘dreams’

b. [-ul] after /p t̪ k d̪ g s v z m ŋ j/

singular	plural	
roːp	roːpul	‘plant’
ket̪	ket̪ul	‘winnowing fans’
maːk	maːkul	‘tree’
mood̪	mood̪ul	‘particular man. . . ’
d̪eg	d̪egul	‘heaps, masses’
kis	kisul	‘fires’
aːv	aːvul	‘fathoms’
gaːz	gaːzul	‘bangle’
d̪em	d̪emul	‘draws on a pipe’
nenjeŋ	nenjeŋul	‘meat’
poj	pojul	‘hearth’

Table 2: Phonologically active class(es) in Kolami.



p		t̪		**t**	k
b		d̪		**d**	g
			tʃ		
			dʒ		
		s			
	v	z			
m		**n̪**			ŋ
		r			
		l			
			j		
i	**iː**			u	uː
e	**eː**			o	oː
		a	**aː**		

condition [-l]

p		**t̪**		t	**k**
b		**d̪**		d	**g**
			tʃ		
			dʒ		
		s			
	v	**z**			
m		n̪			**ŋ**
		r			
		l			
			j		
i	iː			u	uː
e	eː			o	oː
		a	aː		

condition [-ul]

Presumably, no one would propose an innate feature to handle these classes, but there are also less crazy “unnatural” classes which can be described in terms of shared phonetic properties even if they cannot be characterized in terms of traditional distinctive features. One of these occurs in Eastern Cheremis (Sebeok and Ingemann, 1961) (Figure 2). Here, the class of nasals and lateral liquids triggers the lenition of a preceding /d̪/. This class is unnatural in many feature theories because nasals and lateral liquids share no features to the exclusion of all other segments in the inventory (in each theory). Specifically, they share no features to the exclusion of the dental flap. In some feature systems, such as Unified Feature Theory (Clements and Hume, 1995), nasals and laterals are both treated as [+sonorant, –continuant], but so, unfortunately, is the flap.

(6) Not so crazy: Eastern Cheremis[4] (Sebeok and Ingemann, 1961)

a. /d̪/ becomes lenis [đ̪] before /l̪/ and nasals /m n̪ ɲ ŋ/.

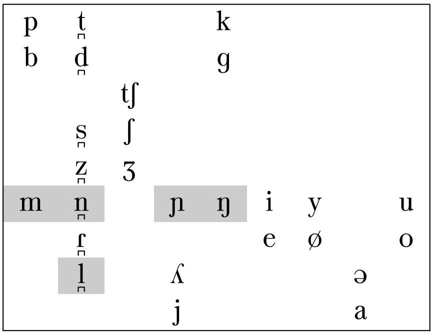
/tud̪l̪an/	[tuđ̪l̪an]	‘to him’
/mod̪maʃ/	[mođ̪maʃ]	‘game’

b. /d̪/ is produced as /d̪/ after nasals.

/ʃən̪d̪as/	[ʃən̪d̪as]	‘to set, put, plant’

c. /d̪/ is reduced to [ð] everywhere else (including before /ɾ̪/).

/tʃod̪ɾ̪a/	[tʃoðɾ̪a]	‘forest’
/lud̪o/	[luðo]	‘duck’



p	t̪			k				
b	d̪			g				
		tʃ						
	s̪	ʃ						
	z̪	ʒ						
m	n̪		ɲ	ŋ	i	y		u
	ɾ̪				e	ø		o
	l̪		ʎ				ə	
			j				a	

Figure 2: Segments which condition /d̪/ lenition (no data for /ʎ/).

This case is less surprising than the Kolami case because nasals and lateral liquids are acoustically similar, both having antiformants generated by side cavities, and so it is not surprising that they pattern together in many languages. Nasals and lateral liquids pattern together in “unnatural” classes in Eastern Cheremis (Sebeok and Ingemann, 1961), Toba (Klein, 2001), and Warlpiri (twice) (Nash, 1986), and many “natural” classes in a wide variety of languages. In most cases, what determines whether or not the class is featurally natural is whether there is another [+sonorant, –continuant] segment in the inventory. The fact that nasals and lateral liquids may pattern together regardless of whether they are a natural class in most feature theories suggest that it may be the shared phonetic property, rather than shared innate features, that is important.

Proposing a feature for this class is more plausible than proposing one for the Kolami classes, and in fact Flemming (2002) proposes an acoustically-defined feature common to laterals and nasals. Adding new features in this way enables classes to be rendered natural while maintaining a distinction between natural and unnatural classes. Such a distinction (between the natural classes that are easily described with features and the idiosyncratic classes which occur even though they are not conducive to innate feature analysis) makes the prediction that there should be some other evidence of a difference between the “privileged” natural classes and the idiosyncratic classes which occur in spite of their incompatibility with the feature set. While we have seen already in Table 1 that featurally unnatural classes are quite widespread, Figure 3 shows how well segregated these classes are from each other.[5] This table shows the distribution of natural and unnatural classes in Unified Feature Theory, which is representative of the three theories (SPE does slightly better and Preliminaries does slightly worse, as seen above in Table 1). Each bar represents a different feature specification (i.e., one or more conjoined features (such as “[+high]” or “[+high, –back]”) for natural classes, or feature disjunction or subtraction (such as “[+high] v [–back]” or “[+high] – [+back, +round]”) for unnatural classes). The height of the column indicates the number of occurrences of that class among the 6077 classes. Light bars (with dark edges) are natural classes, and dark bars are unnatural.[6]

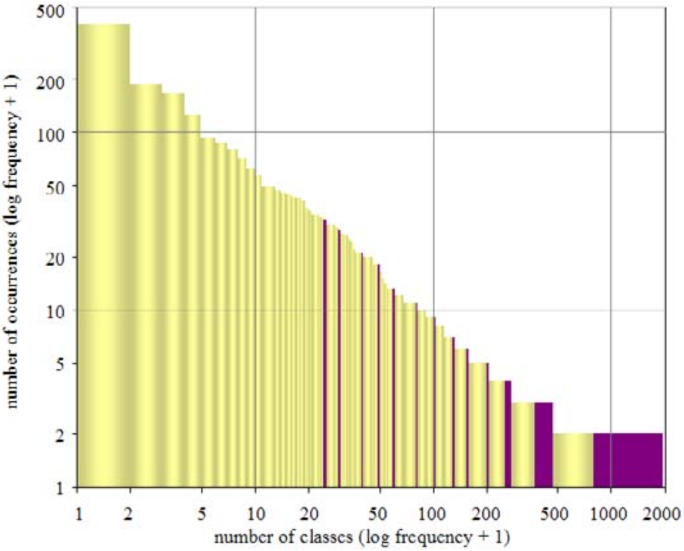


Figure 3: Distribution of frequent and infrequent classes.

Featurally natural and unnatural classes are quite interleaved, with many unnatural classes more frequent than most natural classes, and with the vast majority of natural classes completely unattested. This is another piece of evidence that the distinction between natural and unnatural classes is internal to innate feature theory. The next section further explores a clear similarity between natural and unnatural classes, the fact that both natural and unnatural classes occur among related sound patterns in related languages.

2.2. Unlawful variation

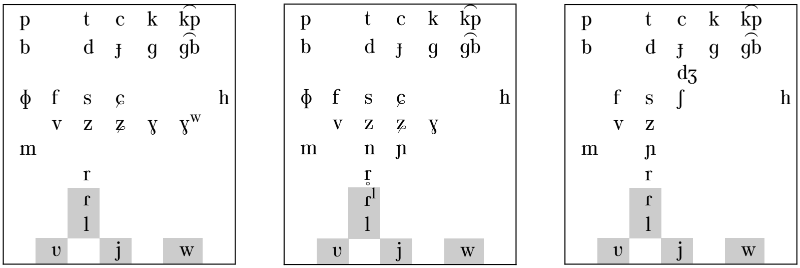
The survey turns up more cases of variation between related languages' interpretation of a shared sound pattern, as in Schaffhausen Swiss German. These cases also suggest that phonological processes generalize in ways that do not correspond to distinctive features. The following examples are from Edoid languages and Russian dialects. See Mielke (2008) for similar cases among Pama-Nyungan and Bantu languages.

In many Edoid languages (Elugbe, 1989), vowels are allophonically nasalized after phonemically nasal consonants, while in many other Edoid languages, the opposite occurs: certain consonants are allophonically nasalized before phonemically nasalized vowels, and the set of targets varies from language to language. Elugbe reconstructs Proto-Edoid with phonemically nasalized consonants, which suggests that the shift to allophonically nasalized consonants is the result of restructuring. In modern Edoid languages with allophonically nasalized consonants, the set of nasalizing consonants varies from language to language. The different classes involve phonetically similar segments, but neither the segments nor the classes are easily related to one another in terms of features. A representative example of consonant nasalization in Edo is shown in (7).

(7) Edo consonant nasalization (Elugbe, 1989: 77, 133-181)[7]

a.	/lɔ̃/	→	[nɔ̃]	'ask'
	/ɹɛ̃/	→	[ɾ̃ɛ̃]	'know'
	/ʋɛ̃/	→	[ʋ̃ɛ̃]	'have'
	/jã/	→	[ɲɛ̃]	'tear apart'
	/wɔ̃/	→	[ŋwɔ̃]	'drink'
b.	/lo/	→	[lo]	'use'
	/a-ɹo/	→	[a-ɹo]	'eye'
	/ʋɛ/	→	[ʋɛ]	'be wide'
	/o-ji/	→	[o-ji]	'thief'
	/wa/	→	[wa]	'you (pl.)'

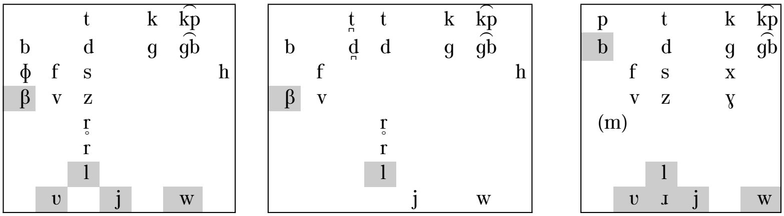
Several Edoid languages with this sound pattern are shown in Figure 4. While the process is similar in all of the languages, the set of involved consonants varies from language to language. These classes include a lateral liquid, tap, and glides in Okpe, Urhobo, and Uvbie (Figure 4a-c)[8], a lateral liquid, glides, and a voiced bilabial fricative in Ehueun (Figure 4d), a lateral liquid and a voiced bilabial fricative in Ukue (Figure 4e), nonnasal sonorants and a voiced bilabial stop in Eruwai (Figure 4f), a lateral liquid, glides, and a velar fricative in Epie (Figure 4g), a lateral liquid, glides, and a glottal fricative in Aoma (4h), and a lateral liquid, an alveolar lateral flap, a glide, and oral stops (which acquire nasal release before nasal vowels) in Edo (4i).



a. Okpe:
[–syl, +son, –nas,
–low, –hi subgl. press.]

b. Urhobo:
[–syl, +son, –nas,
–low, –hi subgl. press.]

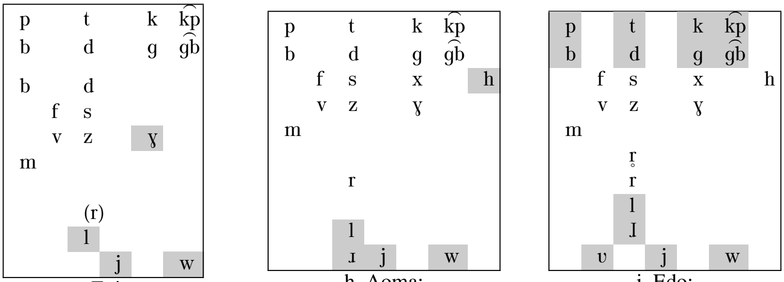
c. Uvbie:
[–syl, +son, –nas,
–low, –hi subgl. press.]



d. Ehueun:
[+voice, +cont, –syl,
–strid, –hi subgl. press.]

e. Ukue:
[+cons, +cont,
–strid, –hi subgl. press.]

f. Eruwa:
[–syl,–nas,+son]
∨ [+voi,–cor,+ant]



g. Epie:
[+cont, –syl, –strid,
(–hi subgl. press.)]

h. Aoma:
[–syl, +son,
–hi subgl. press.]

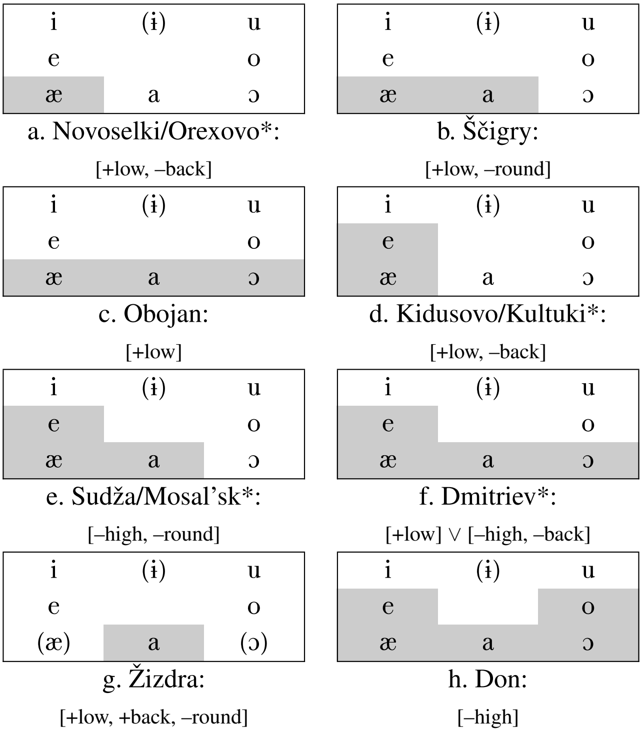
i. Edo:
[+son,+voi,+cont,–syl,
–h.s.p.] ∨ [–nas,–cont]

Figure 4: Nasalizing consonants in Edoid languages.

To summarize, glides, flaps, and /l/ nasalize in most cases. /β/ or /b/ nasalizes in three languages, and /ɣ/ and /h/ each nasalize in one language. All oral stops participate in a similar pattern in one language. Most of these "extra" segments are phonetically similar in some way to the sonorant consonants which nasalize in more languages, but are featurally quite different from them. Two of the languages have passed up perfectly good classes defined by [+voice], [+sonorant], etc., in favor of classes which cannot be specified with a conjunction of features.

A universal feature set predicts that /β/, /ɣ/, /h/, etc. should be systematically included or excluded depending on whether or not features such as [sonorant] are targeted. If they are included as a result of changes in the feature specification for the class, then other fricatives should also be included. Instead, the restructuring process in each language seems to have caused the pattern to be generalized to different sets of phonetically similar segments.

A final example comes from a dissimilation pattern known as *jakan'e* which is found in many dialects of Russian. This case is special because it has been put forth as evidence for innate features, but more recent data has shown that the situation is more complicated. The pattern is one in which unstressed vowels become [i] before certain stressed vowels and [a] before others. The set of stressed vowels which trigger changes to each vowel varies between dialects. Halle (1962) uses examples from several cases to argue that only featurally natural classes appear among the variations on this pattern. However, Culicover (1970) brings data from additional dialects to the question, including one dialect with a featurally unnatural class. The sets of vowels which trigger the change to [i] in each of these varieties are shown in Figure 5. Halle's argument was based on Don, Kidusovo, Novoselki, Obojan, Sudža, Ščigry, and Žizdra, all of which involve featurally natural classes. The additional varieties Culicover adds from Avanesov and Orlova (1965) (Dmitriev, Kultuki, Mosal'sk, and Orexovo) are marked with an asterisk. The features used by Culicover (1970) to define all of these classes are given, and the Dmitriev case requires the disjunction of two feature bundles. This class is problematic for more than just SPE features, because only one of the mid vowels participates, even though all three low vowels (including /ɔ/) do.



a. Novoselki/Orexovo*:
[+low, –back]

b. Ščigry:
[+low, –round]

c. Obojan:
[+low]

d. Kidusovo/Kultuki*:
[+low, –back]

e. Sudža/Mosal’sk*:
[–high, –round]

f. Dmitriev*:
[+low] ∨ [–high, –back]

g. Žizdra:
[+low, +back, –round]

h. Don:
[–high]

Figure 5: Jakan’e [i] triggers in several Russian dialects. * indicates dialects included by Culicover but not by Halle.

Culicover points out that it would be ideal to find that Dmitriev is in contact with two varieties which involve classes whose union is the set of triggers found in Dmitriev (e.g. Obojan and Kidusovo would work), but this is not the case. Rather, Culicover argues for a series of generalizations by which a class (perhaps originally /æ/) is expanded in various dialects, shown in (8).[9] Culicover’s hierarchy of vowel set overgeneralizations is supported by a computer simulation below in Section 4.

(8)

$$\begin{tikzcd}
\text{æ} \arrow[r] & \text{a} \arrow[r] \arrow[d] & \text{e} \arrow[d] \\
 & \text{ɔ} & \text{ɔ} \arrow[d] \\
 & & \text{o}
\end{tikzcd}$$

2.3. Summary

This section has demonstrated that featurally unnatural classes are widespread, that comparable classes in related patterns do not tend to be easily relatable in features, and that natural and unnatural classes can coexist among comparable sound patterns in related language varieties. While innate feature theory generates many of the classes which occur, it cannot account for the many featurally unnatural classes (which may or may not be phonetically natural), and it cannot account for the cases where related languages show both natural and unnatural classes in their versions of the same sound pattern. The next section advances Emergent Feature Theory as a solution.

3. Emergent Feature Theory

Phonetically-defined innate features are just one way to describe classes of phonetically similar segments. There are other ways to describe these classes and to predict common and rare ones, and without innate features, there are reasons to expect natural classes to be frequent anyway. Here I will highlight two opportunities for the development of sound patterns involving phonetically natural classes.

The first opportunity is the inception of a sound change. Phonological patterns can result from a change characterized as phonologization (Hyman 1977), conventionalization, or exaggeration (Janda 1999) of a previously "insignificant" phonetic effect. Phonetic effects involve factors such as physiology and aerodynamics, which do not vary very much between humans. Likewise, languages have similar classes involved in similar phonetically grounded sound patterns which reflect these factors (see also e.g., Dolbey and Hansson (1999)). For example, physiology and inertia dictate that coarticulatory vowel nasalization is likely to be caused by adjacent nasal segments which are produced with velum lowering. Consequently, nasality alternations will tend to involve the familiar class of nasals, regardless of what feature system is used in the synchronic grammar.

The second opportunity is a change in an existing pattern. Classes involved in sound patterns may change over time, and one way for this to happen is for learners to arrive at the "wrong" generalization about a sound pattern, and for the "wrong" generalization to become the prevailing version for a speech community. When this happens, the "wrong" generalization has become "right". The result may be a phonetically natural class that is not necessarily related to the original phonetic basis for the sound pattern.

Most innate feature theories use phonetically-defined features to describe phonetically natural classes. If these features are not innate, then the classes they correctly predict must be attributed instead to the phonetic dimensions the features are grounded in. In Emergent Feature Theory[10] (Mielke, 2008), phonologically active classes can result from generalization to groups of phonetically similar segments. A variety of independently-observable factors such as audition, attention, categorization, aerodynamics, coordination, and social identity may account for many observations about sound patterns. In interpreting the sound patterns confronted during acquisition, a learner constructs abstract features which reflect these factors.

In innate feature theories, sound patterns are built out of features, and phonetics and other "external" factors may be invoked in cases where features are unable to account for observed sound patterns. In Emergent Feature Theory, naturalness in sound patterns is attributed to phonetic pressures and other factors, and phonetics and "external" factors are involved in the development of unnatural *and* natural sound patterns. Features develop in a learner's interpretation of sound patterns confronted in the language being acquired. Phonologically active classes can result from generalization to groups of phonetically similar segments, and phonetics impacts the development of sound patterns directly, rather than being built into an innate feature system. If this is on the right track, then a model of the role of phonetics in language change should be able to replicate hypothesized changes such as Culicover's hierarchy in (8), and this will be seen in the next section.

4. A model of phonological generalization

A simulation of the generalizations hypothesized to occur in the development of sound patterns can help demonstrate that phonetically natural classes may emerge without the help of innate features. Before using phonetic similarity to account for phonological observations, it is necessary to define phonetic similarity. Frisch et al. (1997) offer a phonetic similarity metric based on the number of shared natural classes. Consequently, this requires predetermined features, and it is therefore not applicable to the question of whether classes emerge without them. Instead, what is needed is a similarity metric based on objective measurements of sounds.

Another phonetic similarity metric in Mielke (2005b) is based on acoustic and articulatory information. For this metric, four articulatory dimensions were generated by measuring the size of the lip aperture during the production of 63 crosslinguistically frequent vowels and consonants, the size of the tongue constriction and its location along the palate (both using ultrasound imaging), and the ratio of nasal to oral airflow. Three acoustic dimensions were generated by multidimensional scaling of acoustic distances between the segments, as measured by a dynamic time-warping algorithm applied to matrices of mel-scaled cepstral coefficients based on three waveforms of each segment (See Mielke (2005b) for more details). This similarity metric is used to simulate the emergence of natural classes.

The simulation consists of an inventory of segments, a lexicon, a class of sounds which participate in an arbitrary sound pattern, and a learner, implemented in Python, who can tell how similar sounds are. The learner's task is to figure out what segments are in the class, and given a class of sounds selected from an inventory, the learner tries to infer the class by observing the phonological behavior of words randomly selected from the lexicon. Depending on how parameters are set, the learner may overgeneralize to phonetically-similar segments. This model is shown to mislearn unnatural classes of consonants as natural ones more readily than it mislearns natural classes. For example, the class /p m/ in a language with /b/ in its inventory is readily mislearned as /p b m/, while the learner shows no inclination to mislearn /p b m/ as the less natural /p m/. See Mielke (2005b) for further examples and more detailed methods.

Some adjustable parameters are the width (σ) of the activation function (a normal distribution which activates the target segment and its neighbors) and the value of negative evidence (*NE*), which determines the percentage of the activation value that is subtracted when a segment fails to participate in the sound pattern). Increasing σ increases the effect segments have on their neighbors, and increasing the (negative) value of *NE* reduces the likelihood that positive evidence from neighbors will overwhelm direct evidence that a segment does not participate. Over- and under-generalization is more likely when σ is high, and increasing the value of *NE* makes overgeneralization a more likely outcome, relative to undergeneralization. The learner is able to shift attention to dimensions (among the seven phonetic dimensions) that appear useful for distinguishing its current hypothesis from the rest of the segments in the inventory.

Culicover (1970)'s vowel set generalization hierarchy is a set of five changes, shown again and numbered in (8). The stages correspond to the sets of vowels shown in Figure 5.

(9)

$$\begin{tikzcd}
\text{æ} \arrow[r, "1"] & \text{a} \arrow[r, "2"] \arrow[d, "5"'] & \text{e} \arrow[d, "3"'] \\
& \text{ɔ} & \text{ɔ} \arrow[d, "4"'] \\
& & \text{o}
\end{tikzcd}$$

A simulation with each stage as a starting point should be able to produce the following stage. Specifically, the learner should generalize /æ/ to /æ a/, /æ a/ to both /æ a ɔ/ and /e æ a/, /e æ a/ to /e æ a ɔ/, and /e æ a ɔ/ to /e æ a ɔ o/. Since the learner could conceivably undergeneralize all the way to the empty set and overgeneralize all the way to the entire vowel inventory, it is necessary to set the parameters in a range that allows these particular generalizations to be possible. In general, larger initial classes require more negative values for *NE*, as seen in Table 3 because fewer segments are available to supply negative evidence. The replication of Culicover's generalizations is shown in Table 3. The vowels in the "Input" column correspond to the inputs of each of the five changes in (9).[11]

The simulation was run five times with the same parameters for each input class (five times each with two sets of parameters for /æ a/, which is the input for two different changes). The "Output" column shows how well the simulation fared at replicating the hypothetical changes in each case, and which dialect, if any, the output corresponds to. The change or changes which were replicated are shown in the "Change(s)" column.

Because of the role of phonetic similarity in the simulation, the classes generalize to phonetically similar segments rather than randomly selected segments, and because the model does not involve any innate features, it does not rule out attested classes such as /e æ a ɔ/. In one trial, /æ/ failed to generalize, but in the other four, it either generalized to /æ a/ (change #1) or to /e æ/, which is not in the hierarchy but corresponds to the Kidusovo and Kultuki systems. With one set of parameters, /æ a/ generalized three out of five times to /æ a ɔ/ (change #5), and with another set of parameters, /e æ a/ generalized three out of five times to /e æ a/ (change #2). In the remaining cases, /æ a/ underwent a larger generalization, corresponding to changes #2

Table 3: Simulated Russian vowel set generalization.

NE	σ	Input	Output		Change(s)
-0.10	0.25	æ	æ	(Novoselki/Orexovo)	none
			æ a	(Ščigry)	1
			æ a	(Ščigry)	1
			e æ	(Kidusovo/Kultuki)	*
			e æ	(Kidusovo/Kultuki)	*
-0.15	0.05	æ a	æ a ɔ	(Obojan)	5
			æ a ɔ	(Obojan)	5
			æ a ɔ	(Obojan)	5
			e æ a ɔ	(Dmitriev)	2 & 3
			e æ a ɔ	(Dmitriev)	2 & 3
-0.20	0.50	æ a	e æ a	(Sudža/Mosal'sk)	2
			e æ a	(Sudža/Mosal'sk)	2
			e æ a	(Sudža/Mosal'sk)	2
			e æ a ɔ	(Dmitriev)	2 & 3
			e æ a ɔ	(Dmitriev)	2 & 3
-0.26	0.30	e æ a	e æ a	(Sudža/Mosal'sk)	none
			e æ a	(Sudža/Mosal'sk)	none
			e æ a ɔ	(Dmitriev)	3
			e æ a ɔ	(Dmitriev)	3
			e æ a ɔ	(Dmitriev)	3
-0.35	0.30	e æ a ɔ	e æ a ɔ o	(Don)	4
			e æ a ɔ o	(Don)	4
			e æ a ɔ o	(Don)	4
			e æ a ɔ o	(Don)	4
			i e æ a ɔ o	(unattested)	unattested

and #3. /e æ a/ generalized to /e æ a ɔ/ in all three cases where generalization occurred (change #3), and /e æ a ɔ/ generalized to /e æ a ɔ o/ (change #4) in four out of five cases, and generalized to the unattested /i e æ a ɔ o/ in one case.

In this simulation, Emergent Feature Theory is able to account for the different classes that condition *jakan'e* without relying on innate features or ruling out attested patterns. The simulation readily generalizes to classes such as /e æ a ɔ o/ (nonhigh vowels), which are very natural in innate feature theories, but also generalizes to /e æ a ɔ/, the Dmitriev system, which is problematic for innate features. In reality (as in the simulation), generalization to new classes may or may not occur, evidenced by the fact that even though sound patterns change from time to time, at any given moment they usually stay the same.[12] These simulations support the notion that when a change *does* occur, it is likely to involve phonetically similar segments, accounting for the fact that natural classes are common. Innate features are not necessary to account for this.

5. Conclusions

Although they are frequently marginalized, featurally unnatural classes and patterns of unlawful variation are actually quite numerous, and part of the reality of language. Instead of being treated as exceptions, these natural and widespread phenomena should be embraced and used to inform a more inclusive model of phonology. The insights of work in phonetics, psycholinguistics, sociolinguistics, historical linguistics, and elsewhere help make sense of the types of sound patterns that are observed in language, and in fact many of these insights are codified in innate feature theories. Emergent Feature Theory is simply a means of incorporating these insights into a phonological framework that treats common and rare patterns ("natural" and "unnatural") as natural phenomena, more or less likely responses to the world in which language exists.

Notes

1. This research was made possible by funding from a Presidential Fellowship from The Ohio State University. Thanks also to Mike Armstrong, Peter Culicover, Robin Dodsworth, Beth Hume, Keith Johnson, Brian Joseph, an anonymous reviewer and audience members at VarPhon, NELS 35, WECOL 2004, WCCFL XXIV, and CLS 41.
2. /b/ is less similar to /r/ in some ways than most of the segments which do participate, e.g., most of them are lingual consonants. On the other hand, /b/ is certainly more similar to /r/ than /p^{h}/ is. Because generalization from /r/ likely occurred in more than one step, similarity to /r/ would have been most critical at the stage before other segments began participating, but later it would be similarity to the segments already participating which is relevant, not just similarity to /r/. Perhaps the development of the class in Figure 1e involved an intermediate stage where the class was /r/ and lingual obstruents, and this was further extended to other fricatives and affricates, including /pf f/. Since there are no voiced labiodentals in the language, /p^{h}/ is more similar to the participat-

ing segments than /b/ is at this hypothetical stage. /b/ is also rather similar to /m/, another segment which does not cause /o/ lowering.

3. SPE features (Chomsky and Halle, 1968) are used in this example, but the problem exemplified by 1d and 1e is certainly not unique to this feature theory.
4. A reviewer points out that if this is really a pattern involving various degrees of lenition rather than the three categories suggested by the transcription, then it is not really fair to expect features to account for the middle category. It is worth noting, perhaps, there are other examples of nasal-lateral classes, to the point that a feature has been proposed to account for them (as discussed below).
5. A lack of a distinction between featurally natural and unnatural classes is not problematic for Flemming's approach, which does not assume innate features.
6. Innate feature theories do not rule out the possibility of superficially unnatural classes. In a contrastive inventory, any unnatural group of segments can be characterized as the union of two or more natural classes. If these apparent classes are a product of an innate feature set, the most common unnatural classes should be composed of the most common natural classes. Mielke (2008) shows that this is not the case.
7. For younger speakers, the liquids in 'know' and 'eye' surfaces as [ɪ̃] and [ɹ], respectively.
8. Alternatively, the Urhobo class could be [–syl, +son, –nasal, +voice].
9. This does not account for Žizdra, Kidusovo, or Kultuki. Culicover supports Halle's claim that the Žizdra is probably the result of an Obojan-like pattern being adopted by a dialect with only five vowels. The apparent generalization of /æ/ to /æ e/ in Kidusovo and Kultuki is not particularly alarming either.
10. Emergent Feature Theory is at least partially consistent with and/or inspired by a good deal of work in synchronic and diachronic linguistics, e.g., Martinet (1968); Andersen (1973); Anttila (1977); Anderson (1981); Ohala (1981, 1993); Ladefoged (1984); Lindblom (1984, 1990, 1999); Corina and Sagey (1989); Beddor (1991); Labov (1994, 2001); Port (1996); Steels (1997); Bybee (1998); MacWhinney (1998); Dolbey and Hansson 26 Notes (1999); Buckley (2000); de Boer (2000); Hale and Reiss (2000); Hume and Johnson (2001a); Hyman (2001); Kochetov (2002); Myers (2002); Vaux (2002); Beckman and Pierrehumbert (2003); Hamann (2003); Janda (2003); Janda and Joseph (2003); Pierrehumbert (2003); Pulleyblank (2003); Yu (2003); Blevins (2004); Culicover and Nowak (2004); Hume (2004); Wedel (2004).
11. The Russian results do not appear to be a peculiarity caused by specific choices of dimensions in the similarity model. When scaled to two dimensions, the vowel similarity model resembles an acoustic vowel space, and traditional dimensions of vowel height and backness are robustly present in articulation and acoustics, particularly the oral constriction size and location dimensions and the acoustic dimension interpretable roughly as grave/acute.

12. Because the simulation is concerned with the outcomes of change, enough repetitions are performed so that a change to a new class is likely. The vast majority of individual steps result in no change.

References

Andersen, Henning
1973 Abductive and deductive change. *Language* 49: 765-793.
Anderson, Stephen R.
1981 Why phonology isn't "natural". *Linguistic Inquiry* 12: 493-539.
Anttila, Raimo
1977 *Analogy*. The Hague: Mouton.
Avanesov, R.I., and V.G. Orlova
1965 *Russkaja dialektologija*. Moskva: Izdatel'stvo Mauka.
Beckman, Mary, and Janet Pierrehumbert
2003 Interpreting 'phonetic interpretation' over the lexicon. In *Papers in Laboratory Phonology VI*, J. Local, R. Ogden, and R. Temple (ed.), 13-37. Cambridge: Cambridge University Press.
Beddor, Patrice Speeter
1991 Predicting the structure of phonological systems. *Phonetica* 48: 83-107.
Blevins, Juliette
2004 *Evolutionary phonology*. Cambridge: Cambridge University Press.
de Boer, Bart
2000 *The origins of vowel systems*. Oxford: Oxford University Press.
Buckley, Eugene
2000 What should phonology explain? Handout from SUNY Buffalo linguistics colloquium.
Bybee, Joan
1998 Usage-based phonology. In Darnell et al. (1998), 211-242.
Chomsky, Noam, and Morris Halle
1968 *The Sound Pattern of English*. Cambridge, Mass.: MIT Press.
Clements, G.N., and Elizabeth V. Hume
1995 The internal organization of speech sounds. In *The Handbook of Phonological Theory*, John Goldsmith (ed.), 245-306. Cambridge Mass.: Blackwell.
Corina, David P., and Elizabeth Sagey
1989 Are phonological hierarchies universal? Evidence from American Sign Language. In *Proceedings of the Sixth Eastern States Conference on Linguistics*, Ken de Jong and Yongkyoon No (ed.), 73-83.
Culicover, Peter
1970 Akan'e, Jakan'e, Ikan'e, Jokan'e. Ms.

Culicover, Peter W., and Andrzej Nowak
2004 *Dynamical grammar*. Oxford: Oxford University Press.

Darnell, M., E. Moravcsik, F. Newmeyer, M. Noonan, and K. Wheatley (eds.)
1998 *Functionalism and Formalism in Linguistics Volume 1: General Papers*. Amsterdam: John Benjamins.

Dolbey, A.E., and G.Ó. Hansson
1999 The source of naturalness in synchronic phonology. In *CLS 35, Vol. 1*, S. Billings, J. Boyle, and A. Griffith (ed.), Chicago: CLS.

Elugbe, Ben Ohiomamhe
1989 *Comparative Edoid: Phonology and lexicon*. Port: University of Port Harcourt Press.

Emeneau, M. B.
1961 *Kolami: A Dravidian language*. Annamalainagar: Annamalai University.

Flemming, Edward S.
2002 *Auditory representations in phonology*. New York: Routledge.

Frisch, Stefan, Janet Pierrehumbert, and Michael Broe
1997 Similarity and phonotactics in Arabic. Indiana University ms., ROA-223.

Hale, Mark, and Charles Reiss
2000 Phonology as cognition. In *Phonological knowledge: Conceptual and empirical foundations*, N. Burton-Roberts, Philip Carr, and Gerard Docherty (ed.), 161-184. Oxford: Oxford University Press.

Halle, Morris
1962 Akan'e, the treatment of unstressed nondiffuse vowels in Southern Great Russian dialects. In *Symbolae linguisticae in honorem Georgii Kurylowicz*, Adam Heinz et al. (eds.), 103-109. Wroclaw: Polska Akademia Nauk.

Hamann, Silke
2003 *The phonetics and phonology of retroflexes*. Ph. D. thesis, University Utrecht.

Hume, Elizabeth V.
2004 The indeterminacy/attestation model of metathesis. *Language* 80: 203-237.

Hume, Elizabeth V., and Keith Johnson
2001a A model of the interplay of speech perception and phonology. In Hume and Johnson (2001b), 3-26.

Hume, Elizabeth V., and Keith Johnson, ed.
2001b *The role of speech perception in phonology*. New York: Academic Press.

Hyman, Larry M.
1977 Phonologization. In *Linguistic studies presented to Joseph H. Greenberg*, A. Juilland (ed.), 407-418. Saratoga: Anna Libri.

Hyman, Larry M.
2001 The limits of phonetic determinism in phonology: *NC revisited. In Hume and Johnson (2001b).

Jakobson, Roman, C. Gunnar M. Fant, and Morris Halle
1954 *Preliminaries to speech analysis: the distinctive features and their correlates*. Cambridge, Mass.: MIT Press.

Janda, Richard D.
1999 Accounts of phonemic split have been exaggerated - but not enough. In *Proceedings of The XIVth International Congress of Phonetics Science* John J. Ohala et al. (eds.), 329-332.

Janda, Richard D.
2003 "Phonologization" as the start of dephoneticization – or, on sound-change and its aftermath: Of extension, generalization, lexicalization, and morphologization. In Joseph and Janda (2003), 401-422.

Janda, Richard D., and Brian D. Joseph
2001 Reconsidering the canons of sound-change: Towards a big bang theory. Ohio State University ms.

Janda, Richard D., and Brian D. Joseph
2003 On language, change, and language change – or, of history, linguistics, and historical linguistics. In Joseph and Janda (2003), 3-180.

Joseph, Brian D., and Richard D. Janda, (eds.)
2003 *The Handbook of Historical Linguistics*. Oxford, UK: Blackwell.

Keel, William
1982 *Atomic phonology and phonological variation*. Tübingen: Gunter Narr.

Klein, Harriet E. Manelis
2001 *Toba*. München: Lincom Europa.

Kochetov, Alexei
2002 *Production, perception, and emergent phonotactic patterns: A case of contrastive palatalization*. New York: Routledge.

Labov, William
1994 *Principles of linguistic change: Internal factors*. Oxford: Blackwell.

Labov, William
2001 *Principles of linguistic change: Social factors*. Oxford: Blackwell.

Ladefoged, Peter
1984 Out of chaos comes order; physical, biological, and structural patterns in phonetics. In *Proceedings of ICPhS 10*, M.P.R. Van den Broecke and A. Cohen (ed.), 83-95. Dordrecht: Foris Publications.

Lindblom, Björn
1984 Can the models of evolutionary biology be applied to phonetic problems? In *Proceedings of ICPhS 10*, M.P.R. Van den Broecke and A. Cohen (ed.), 67-81. Dordrecht: Foris Publications.

Lindblom, Björn
1990 Explaining phonetic variation: A sketch of the H&H theory. In *Speech production and speech modeling*, William J. Hardcastle and Alain Marchal (ed.), 403-439. Dordrecht: Kluwer.

Lindblom, Björn
1999 Emergent phonology. University of Stockholm/University of Texas at Austin ms.

MacWhinney, Brian
1998 Emergent language. In Darnell et al. (1998), 361-386.

Martinet, Andre
1968 Phonetics and linguistic evolution. In *Manual of Phonetics*, Bertil Malmberg (ed.). Amsterdam: North-Holland Publishing Company.

Mielke, Jeff
2005a Ambivalence and ambiguity in laterals and nasals. *Phonology* 22.

Mielke, Jeff
2005b Modeling distinctive feature emergence. In *WCCFL 24*.

Mielke, Jeff
2008 *The emergence of distinctive features.* Oxford: Oxford University Press.

Myers, Scott
2002 Gaps in factorial typology: The case of voicing in consonant clusters. University of Texas at Austin ms., ROA-509.

Nash, David
1986 *Topics in Warlpiri grammar*. New York: Garland Publishing, Inc.

Ohala, John J.
1981 The listener as a source of sound change. In *CLS 17: papers from the parasession on language and behavior*, C.S. Masek, R.A. Hendrik, and M.F. Miller (ed.), 178-203. Chicago: CLS.

Ohala, John J.
1993 Sound change as nature's speech perception experiment. *Speech Communication* 13: 155-161.

Pierrehumbert, Janet
2003 Probabilistic phonology: Discriminability and robustness. In *Probabilistic linguistics*, R. Bod, J. Hay, and S. Jannedy (eds.), 177-228. Cambridge, Mass.: MIT Press.

Port, Robert F.
1996 The discreteness of phonetic elements and formal linguistics: response to A. Manaster Ramer. *Journal of Phonetics* 24: 491-511.

Pulleyblank, Douglas
2003 Covert feature effects. In *Proceedings of WCCFL 22*, Gina Garding and Mimu Tsujimura (eds.), 398-422. Somerville, Mass.: Cascadilla Press.

Sebeok, Thomas A., and Frances J. Ingemann
1961 *An Eastern Cheremis manual*. Bloomington/The Hague: Mouton and Co./ Indiana University.

Steels, Luc
1997 The synthetic modelling of language origins. *Evolution of Communication* 1: 1-34.

Vaux, Bert
2002 Consonant epenthesis and the problem of unnatural phonology. Paper presented at Yale University Linguistics Colloquium.

Wedel, Andrew
2004 *Self-organization and categorical behavior in phonology.* Ph. D. thesis, UC Santa Cruz.

Yu, Alan C. L.
2003 *The morphology and phonology of infixation.* Ph. D. thesis, UC Berkeley.

Phonetic variation and gestural specification: Production of Russian consonants

Alexei Kochetov

1. Introduction

Human speech is known to exhibit variability at multiple levels. The focus of many phonetic and sociolinguistic theories (e.g. Keating 1990; Lindblom 1990; Labov 2001) has been on identifying and explaining various sources of *phonetic variability* – variable physical realization of categorical, symbolic phonological structure. One influential approach to explaining phonetic variability has been offered by the framework of *Articulatory Phonology* (AP; Browman and Goldstein 1989, 1990, 1992; Goldstein, Byrd, and Saltzman 2006). The key proposal of AP is that much of the observed variation in speech follows from combinatorial dynamics of basic units of speech, *articulatory gestures*.

The gestures are defined as linguistically-significant movements of articulators, which are both physical and cognitive in nature: units of action and units of contrast. AP assumes a *Gestural Linguistic Model* which consists of a universal set of gestures, each specified for one or two tract variables (Browman and Goldstein 1989). For example, the Tongue Tip gesture is specified for numerical values of the Tongue Tip *constriction degree* (TTCD) and the Tongue Tip *constriction location* (TTCL). These values are assumed to be language-particular, while falling within certain ranges that are defined by gestural descriptors. For TTCD, for example, the descriptors are 'closed' (for stops), 'critical' (for fricatives), 'narrow' (for approximants), etc. For TTCL the descriptors are 'dental', 'alveolar', 'post-alveolar', etc. Thus, a partial gestural specification for alveolar oral and nasal stops /t d n/ would be Tongue Tip [CD: closed, CL: alveolar]. This specification defines the 'rest position': the Tongue Tip is instructed to form a complete constriction at the alveolar ridge.

Each tract variable is additionally specified for *stiffness* and *damping ratio*. The first one is a dynamic parameter ultimately responsible for the duration of the gesture: the greater the stiffness, the sooner the rest position is achieved, and the thus the shorter the resulting gesture will be. The damping ratio determines whether the tract variable overshoots the rest position value, undershoots it, or fails to approach it. In the original implementation of the Gestural Linguistic Model only two stiffness values and one damping ratio values were employed: lower stiffness for vowel gestures, higher stiffness for consonant gestures, and critical damping (no overshoot) for all gestures (Browman and Goldstein 1990: 305–306). The rest position, stiffness, and damping ratio values altogether define a task dynamic regime of a gesture, modeled as a damped mass spring equation ('point attractor') (Browman and Goldstein 1990; Saltzman and Byrd 2000). Individual gestures, 'atoms', combine with each other to form larger combinations, 'molecules' (segments, clusters, syllable positions, etc.). In these combinations, timing of gestures is assumed to be governed by 'phasing' rules: a certain point in the trajectory of one gesture is phased with respect to a certain point in the trajectory of another gesture (Browman and Goldstein 1990: 308–310). Gestures can be phased within what is considered to be a single segment, or across different segments (Byrd 1996b: 159–162).

Since its inception, AP has provided researchers with a powerful tool to gain insight in a number of phonetic/phonological phenomena, such as consonant assimilation and deletion, allophonic variation, syllable structure, and recently prosody, phonotactics, and speech errors (see Byrd 2003 for review). The original Gestural Linguistic Model, based on a number of "simplifying assumptions and hypothesized generalizations" (Browman and Goldstein 1990: 305), was put to test in a number of empirical studies (e.g. Zsiga 1995; Byrd 1996a; Gick 2003, among many others). While having provided substantial support for the model, many studies have raised additional issues. Particularly, questions still remain about how the full extent of within-speaker, within-language, and cross-linguistic phonetic variation can be successfully captured relying on task-dynamics and the current representational assumptions of AP (cf. Byrd 2003; Kochetov 2006a).

The primary goal of this paper is to identify possible sources of phonetic variation by testing several specific predictions made by the Gestural Linguistic Model of AP. This is done by a thorough magnetic articulometer investigation of within-speaker phonetic variation, examining the production of 12 Russian consonants in three different contexts (over 450 tokens).[1] The secondary goal of the paper is to provide initial empirical data for an implementation of a language-particular (Russian) version of the Gestural Linguistic Model with the purpose of further testing our understanding of cross-linguistic speech production.

2. Gestural specification and phonetic variation in Russian consonants

2.1. Assumptions

The 12 Russian consonants examined in this paper are plain/palatalized labial and coronal stops, nasals, and fricatives. We will assume that these consonants are stable combinations of several invariant gestures, as shown in Table 1. Gestural descriptions are based on Browman and Goldstein (1989); descriptor values are estimated from phonetic literature on Russian (Bolla 1981; Kochetov 2006b).

Table 1: Combinations of articulatory gestures resulting in Russian plain/palatalized labial and coronal stops, nasals, and fricatives (preliminary specifications).

Cons.	Primary constriction gesture	Secondary constriction gesture	Other gestures
/p/	Lips [closed, labial]	TB [narrow, velar]	Glottis [wide]
/m/	Lips [closed, labial]	TB [narrow, velar]	Velum [open]
/pʲ/	Lips [closed, labial]	TB [narrow, palatal]	Glottis [wide]
/mʲ/	Lips [closed, labial]	TB [narrow, palatal]	Velum [open]
/f/	Lips [critical, dental]	TB [narrow, velar]	Glottis [wide]
/fʲ/	Lips [critical, dental]	TB [narrow, palatal]	Glottis [wide]
/t/	TT [closed, dental]	*n/a*	Glottis [wide]
/n/	TT [closed, dental]	*n/a*	Velum [open]
/tʲ/	TT [closed, dental]	TB [narrow, palatal]	Glottis [wide]
/nʲ/	TT [closed, dental]	*n/a*	Velum [open]
/s/	TT [critical, alveolar]	*n/a*	Glottis [wide]
/sʲ/	TT [critical, alveolar]	TB [narrow, palatal]	Glottis [wide]

Note: TT = Tongue Tip; TB = Tongue Body.

The bilabials /p m pʲ mʲ/ and labio-dentals /f fʲ/ are specified here for the gestures Lips [closed, labial] and Lips [critical, dental] respectively. The dental/alveolar (coronals) stops/nasals and fricatives are assumed to be specified for the TT [closed, dental] and TT [critical, alveolar] respectively. All palatalized consonants are specified for the TB [narrow, palatal]; ‘plain’ labials, which are velarized, are specified here for the TB [narrow, velar]; plain coronals are not specified for the TB gesture. All voiceless stops are specified for Glottis [wide]; all nasals are specified for Velum [open]. The latter two gestures will not be investigated in the study. We will further assume that all the tract variables in Table 1 have the same values for stiffness and damping (cf. Browman and Goldstein 1990: 305–306).

Previous work has shown that the TB gesture for syllable-final palatalized labial stops in Russian is phased at the achievement of, or simultaneously with, the Lip gesture (Kochetov 2006a). We will assume this phasing of TB for all relevant consonants: palatalized labials (TB-Lips), plain (velarized) labials (TB-Lips), and palatalized coronals (TB-TT; with possible modifications described below).

2.2. Predictions

The central assumption of AP that gestures are 'atoms' of speech, can be interpreted that gestural specifications are invariant regardless of whether a given gesture is part of one or another segment. For example, presumably the Lip specifications for /p/ and /m/ or the TB specifications for /fʲ/ and /sʲ/ have the same values, even though these can be seen as different higher-level units, segments (cf. Byrd 1996b: 159–162 on segments being epiphenomenal). This hypothesis will be referred to in the paper as the *gestural invariance hypothesis*. The hypothesis can be tested when a gesture does not overlap (is not activated simultaneously) with other gestures that are mechanically coupled with it. For example, the Lips and the TB of /fʲ/ are mechanically uncoupled, while the TT and TB of /sʲ/ are coupled.

Mechanically coupled gestures are expected to interfere with each other, particularly when they have conflicting rest positions (e.g. TT [dental] and [alveolar] or [closed] and [critical]) – the phenomenon called gestural 'blending' (Browman and Goldstein 1992: 160–161). Thus, a simultaneous activation of TT [dental] and TB [palatal] may result in a more posterior constriction of the TT and in a more anterior constriction of the TB (e.g. within the segment /sʲ/). The phasing of the two gestures may also be affected. A simultaneous activation of Lips [critical, dental] and Lips [closed, labial] may result in trajectories intermediate between the specified [critical] and [closed] values, as well as between the [dental] and [labial] values (e.g. across segments in /f#p/). This hypothesis will be referred to as the *gestural blending hypothesis*.

Finally, syllable-final gestures are known to reduce in magnitude and duration compared to the same gestures syllable-initially (Browman and Goldstein 1995; Krakow 1999; Gick 2003). The TT gesture of stops and nasals (TT [closed] of /t n/) seems to be particularly susceptible to reduction (also the TB [narrow] of /pʲ/; Kochetov 2006a), especially, if it is followed by another oral gesture (the Lips or the Tongue Body; Browman and Goldstein 1995; Byrd 1996a). This hypothesis will be referred to as the *gestural reduction hypothesis*.

All three hypotheses will be tested in the following study. The focus of Experiment 1 will be on both *gestural invariance* and (within-segment) *gestural blending hypotheses*. The focus of Experiment 2 will be on the (across-segment) *gestural blending hypothesis*. The focus of Experiment 3 will be on the *gestural reduction hypothesis*.

3. Experiment 1

The goal of this experiment is to test the *gestural invariance* and the (within-segment) *gestural blending* by examining single syllable-final labial and coronal stops, nasals, and fricatives /p pʲ m mʲ f fʲ t tʲ n nʲ s sʲ/.

3.1. Method

Data for the experiment were collected using the EMMA (Electromagnetic Midsagittal Articulometer: Perkell et al. 1992) system at Haskins Laboratories. The data were collected from one male speaker of standard Russian, the author. The stimuli included nonsense and real Russian words with 12 single word-final consonants (Table 2). Nonsense utterances and corresponding real word utterances had the same target the same immediate environments and stress pattern (e.g. [ˈtap ˈapɨ] or [ˈɡrap ˈadɨ]).

Real word utterances were either verb (imperative/infinitive) + noun (accusative sg.) or noun (nominative sg.) + noun (genitive sg.) combinations. All stimuli, embedded in a carrier phrase [ˈɛtʌ ___ ʌˈpʲatʲ] ‘This is ___ again’, were presented in alternating blocks of nonsense and real word utterances. Five tokens of each nonsense utterance and four tokens of each real word utterance were collected, yielding a total of 108 tokens (9 tokens per each consonant).

Table 2: Stimuli used in Experiment 1.

Consonant		Nonsense word stimuli	Real word stimuli
a. Labial	/p/	ta**p** apɨ	kra**p** adɨ
	/pʲ/	ta**pʲ** apɨ	gra**pʲ** adu
	/m/	ta**m** apɨ	ada**m** adɨ
	/mʲ/	ta**mʲ** apɨ	dina**mʲ** adu
	/f/	ta**f** apɨ	ʐɨra**f** adɨ
	/fʲ/	ta**fʲ** apɨ	patra**fʲ** adu
b. Coronal	/t/	ta**t** apɨ	bra**t** adɨ
	/tʲ/	ta**tʲ** apɨ	bra**tʲ** adu
	/n/	ta**n** apɨ	kra**n** adɨ
	/nʲ/	ta**nʲ** apɨ	gra**nʲ** adɨ
	/s/	ta**s** apɨ	akra**s** adɨ
	/sʲ/	ta**sʲ** apɨ	ukra**sʲ** adu

Receivers for the articulometer were placed at the following midsagittal points: *Upper Lip* and *Lower Lip* (UL and LL), lower incisors (as an estimate of *Jaw* movement), and four points on the tongue, among which the *Tongue Tip* (TT; about 5 mm from the actual tip) and *Tongue Body* (TB) will be relevant. The placement of the receivers is shown in the spatial display in Figure 1 (the top panel).

The movement data collection and preparation were the same as in Kochetov (2006a). The analysis included the following measurements.

Gestural magnitude measurements (see Figure 1, the top panel):

- The vertical and horizontal position of the Lower Lip (LLx and LLy) and the Upper Lip (ULx and ULy) was measured (in mm) for labial consonants at the LLy maximum and at the ULy minimum respectively; Lip Aperture (LA) was calculated as the vertical distance between the UL and LL (in mm).

- The position of the Tongue Tip (TTx and TTy) was measured for coronal consonants at the TTy maximum.

- The position of the Tongue Body (TBx and TBy) was measured at the TBy maximum for palatalized consonants and the release of the primary gesture (LA or TT) for plain consonants.

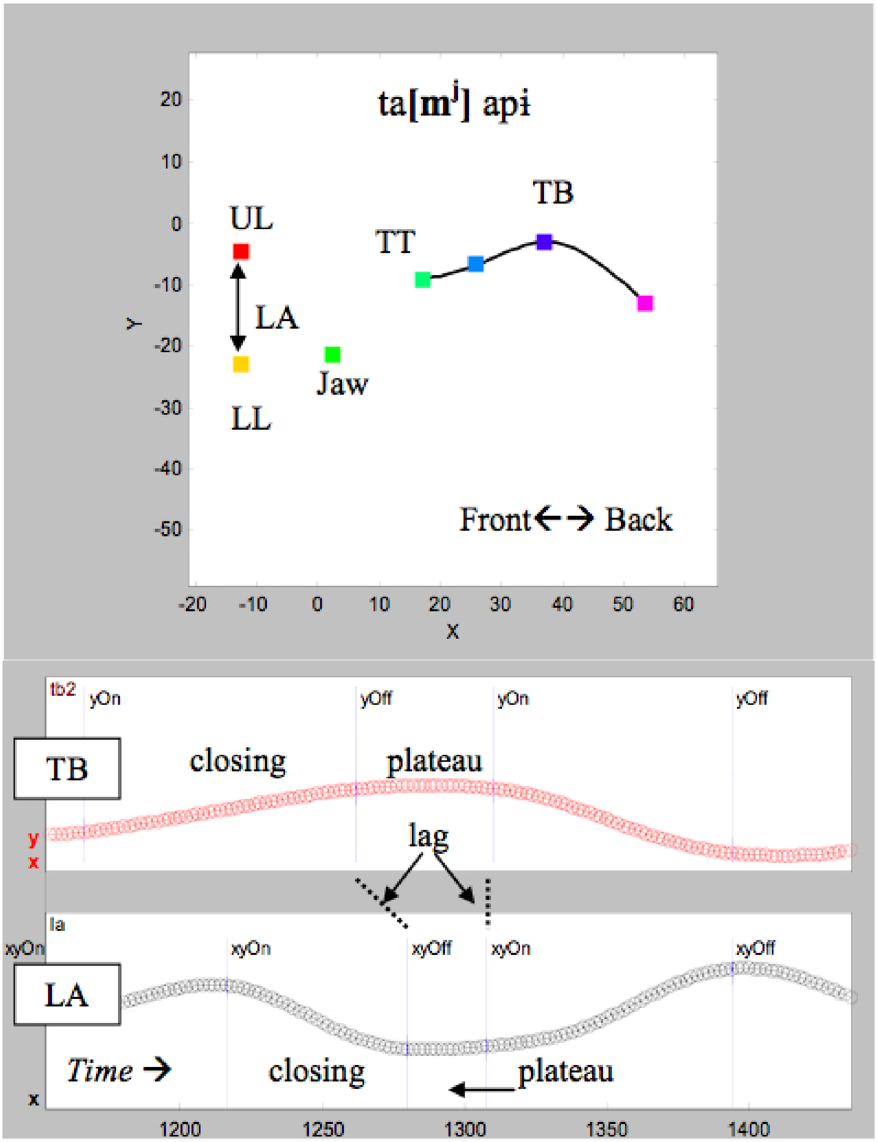


Figure 1: A sample sagittal display of receiver positions (in mm) during the production of /mʲ/ in the context [ta__ apɨ] (at the midpoint of TB plateau) at the top and vertical trajectories of the Tongue Body (TB) and Lip Aperture (in ms) at the bottom.

Gestural duration and timing measurements (see Figure 1, the bottom panel):

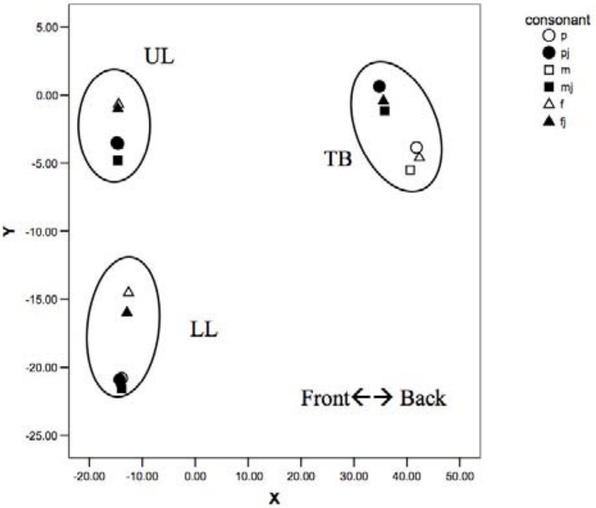
- The duration of the gestures of the LA, TT, and TB: from the onset of the movement towards a constriction to its offset (*closing duration*) and from the previous point to the onset of the movement away from this constriction (*plateau duration*). For the TB, the overall duration (closing + plateau) was also calculated.
- Achievement lag and release lag – periods of time between the achievement or release of two constrictions (TB-LA or TB-TT) – were measured for palatalized consonants.

Gestural magnitude measurements were performed for both nonsense and real word utterances. Gestural duration and timing measurements were performed for nonsense utterances only.[2] An ANOVA with factors Manner (stop, nasal, and fricative) and Palatalization (plain and palatalized) was performed to determine significant effects and interactions. Tukey HSD post hoc tests were performed to investigate relative effects of levels within the factor of Manner. Due to space limitations, statistical results are not presented. Unless stated otherwise, all differences reported in the paper are significant ($p < .05$).

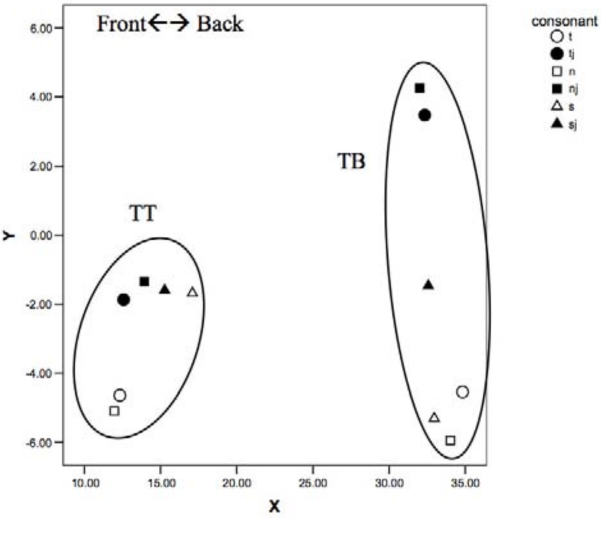
3.2. Results and discussion

Figure 2 (a) plots mean Lower Lip (LL), Upper Lip (UL), and the Tongue Body (TB) magnitude values for labial consonants. Figure 2 (b) plots mean Tongue Tip (TT) and Tongue Body values for coronal consonants. Figure 3 shows mean closing and plateau duration values for labial consonants (a) and coronal consonants (b). For plain consonants, only the primary gesture values are shown (LA or TT); for palatalized consonants, these are accompanied by the TB gesture values. Mean lag measurements are shown by dotted lines.

The results for labial consonants partly support the *gestural invariance hypothesis*. Lip gesture magnitude values (LA, LLy, LLx, and ULy) tend to cluster in two distinct areas, corresponding to two tract variables: Lips [closed, labial] of stops and nasals /p pʲ m mʲ/, and Lips [critical, dental] of fricatives /f fʲ/ (Figure 2a).[3] This tract variable difference is also manifested in different stiffness values: a slower movement towards the fricative target Lips [critical, dental] of /f fʲ/ (Figure 3a) (cf. Browman and Goldstein 1990: 306).

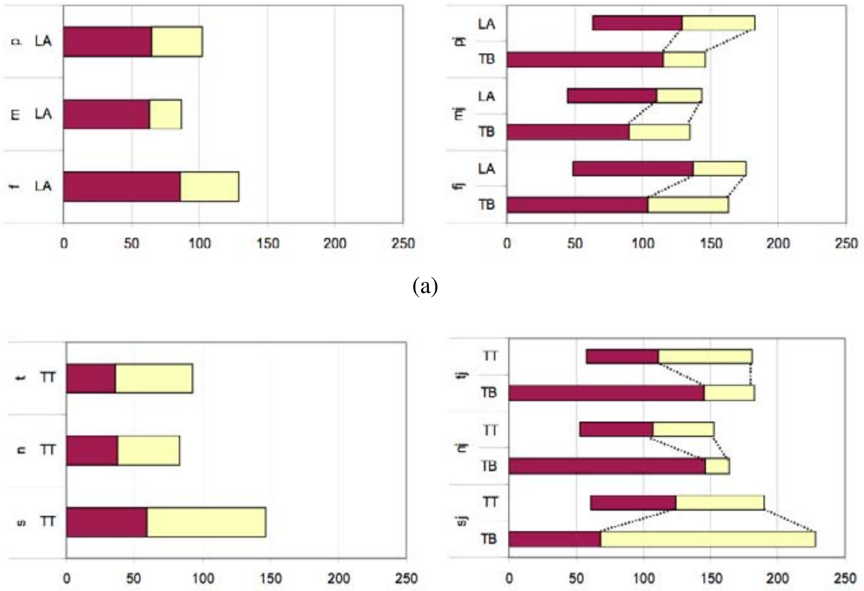


(a)



(b)

Figure 2: (a) Mean magnitude values (in mm) of labial consonants [p p^j m m^j f f^j]: Upper Lip (UL), Lower Lip (LL), and Tongue Body (TB). (b) Mean magnitude values of coronal consonants [t t^j n n^j s s^j]: Tongue Tip (TT) and Tongue Body (TB).



(a)

(b)

Figure 3: (a) Mean Lip Aperture (LA) and Tongue Body (TB) duration values - closing and plateau - (in ms) of labial consonants [p p^j m m^j f f^j]; (b) mean Tongue Tip (TT) and Tongue Body (TB) duration of coronal consonants [t t^j n n^j s s^j].

Contrary to the hypothesis, however, the stiffness specifications differ for nasals /m m^j/ and stops /p p^j/ (a shorter plateau for the nasals; Figure 3a). As expected, the TB specifications for the palatalized and plain labials are very different: the more front and higher TB [narrow, palatal] of /p^j m^j/ and the more back and lower TB [narrow, velar] of /p m/ (Figure 2a). Surprisingly, however, the TBCD ([narrow]) specification for /p^j/ is somewhat different from /m^j/ (a narrower constriction for the stop; neither are different from /f^j/; Figure 2a). The stiffness of the TB gesture also appears to vary depending on whether it is phased with Lips [closed, labial] for /p^j m^j/ (shorter overall duration) or with the Lips [critical, dental] for /f^j/ (longer overall duration). Despite these differences, the phasing of the TB and LA is the same for all palatalized labial consonants: the peak of the TB is phased at the achievement of the LA (negative achievement and release lag values, Figure 3a; cf. Kochetov 2006a).

The results for coronal consonants also provide some support for the *gestural invariance hypothesis*. The specifications TT [closed, dental] and TT [critical, alveolar] result in different TT position values for plain stops/nasals (presumably a closure at the upper teeth for /t tʲ n nʲ/) and fricatives (presumably a narrow opening between the tongue lamina and the alveolar ridge for /s sʲ/; Figure 2b). The two gestures also appear to have different stiffness specifications: a higher stiffness for TT [closed, dental] (a shorter closing movement for /t tʲ n nʲ/) than for TT [critical, alveolar] (a longer closing movement for /s sʲ/; Figure 3b). Again, however, stops and nasals appear to have different stiffness specifications resulting in different plateau durations (shorter for /n nʲ/ than for /t tʲ/).

As expected, the TB [narrow, palatal] gesture of palatalized coronals /tʲ nʲ sʲ/ results in a higher and (for stops and nasals) more front position of the TB than for plain coronals /t n s/, which are presumably unspecified for the gesture (Figure 2b). Interestingly, the TB [narrow, palatal] for coronals /tʲ nʲ sʲ/ is higher and more front than for labials /pʲ mʲ fʲ/ (cf. Figure 2a), a possible consequence of its blending with the TT. The constriction degree and stiffness of the TB gesture differ depending on whether these variables are coordinated with the TT [closed, dental] (a higher TB of a shorter duration for /tʲ nʲ/) or TT [critical, alveolar] (a lower TB of a longer duration for /sʲ/) (Figures 2b and 3b). So does the TB-TT phasing: the TB of the stops/nasals is phased at the release of the TT (positive achievement and release lag values for /tʲ nʲ/); the TB of the fricative is phased roughly simultaneously with the constriction of TT (negative achievement lag and positive release lag values for /sʲ/; Figure 3b). Note that both phasing patterns are very different from the phasing in palatalized labials /pʲ mʲ fʲ/ (cf. Figure 3a). The different timing for the coronals can be attributed to the mechanical coupling of TB and TT. For /tʲ/, this TB-TT timing is possibly related to its characteristic high frequency burst/frication (Kochetov 2006b). The durational characteristics of the TB of palatalized coronals /tʲ nʲ sʲ/ are also different from those of the TB of palatalized labials /pʲ mʲ fʲ/ (Figure 3ab), suggesting different stiffness specifications (or a possible effect of blending). The more back and lower position of the TB for plain labials /p m f/ than for plain coronals /t n s/ reflects the TB [narrow, velar] specification (velarization) for the labials and no such specification for the coronals (Figure 2ab; note scale differences).

The results for coronal consonants also provide some support for the *gestural blending hypothesis*. The simultaneous activation of the TT [closed, dental] and TB [narrow, palatal] for palatalized stops and nasals /tʲ nʲ/ results in some intermediate CD and CL values: a higher and more posterior TT constriction (presumably a closure at the alveolar ridge) compared to plain coronals /t n/, and a higher and more anterior TB constriction compared to palatalized labials /p^j m^j/ (Figure 2ab). The resulting trajectories thus exhibit a blending of tract variables. The assumed specification, however, does not necessarily predict the observed degrees of blending (e.g. the alveolar rather than post-alveolar articulation for /t^j n^j/) or lack of such blending (for /s^j/).

4. Experiment 2

The goal of this experiment is to test the (across-segment) *gestural blending hypothesis* by examining syllable-final labial and coronal stops, nasals, and fricatives /p p^j m m^j f f^j t t^j n n^j s s^j/ before syllable-initial homorganic stops /p p^j t t^j/.

4.1. Method

Stimuli for the experiment are shown in Table 3. Data were collected from the same speaker, using the same procedure as in Experiment 1. A total of 180 tokens were collected (20 nonsense utterances * 5 repetitions + 20 real word utterances * 4 repetitions), or 9 tokens per each sequence.

The analysis included measurements of the Lip Aperture (LA), Tongue Tip (TT), and Tongue Body (TB) magnitude and duration, achievement lag, and release lag, as defined in Experiment 1. (Note that homorganic sequences in the articulatory display were rendered having single LA or TT constrictions. Similarly, C^j#C^j sequences had a single TB constriction.) For C#C sequences, TB magnitude was measured at the midpoint of the acoustic closure of the sequence (for stop-initial sequences only); no TB duration measurements were possible. As before, duration and timing measurements were limited to nonsense utterances. An ANOVA with the factors C1 Manner (stop, nasal, and fricative), Palatalization (C#C, C^j#C, C#C^j, and C^j#C^j), and Place (labial and coronal) was performed to determine significant effects and interactions, followed up by Tukey HSD post hoc tests. Another ANOVA with the factor Context was performed to compare the homorganic sequence contexts C#C and C^j#C with the single consonant condition (C#a and C^j#a).

Table 3: Stimuli used in Experiment 2.

Consonant sequence		Nonsense word	Real word
Place and Manner	Palatalization	stimuli	stimuli
labial stop-stop	C#C	ta**p** **p**apɨ	kra**p** **p**adaja
	Cʲ#C	ta**pʲ** **p**apɨ	gra**pʲ** **p**adaja
	C#Cʲ	ta**p** **pʲ**apɨ	kra**p** **pʲ**atava
	Cʲ#Cʲ	ta**pʲ** **pʲ**apɨ	gra**pʲ** **pʲ**atava
labial nasal-stop	C#C	ta**m** **p**apɨ	ada**m** **p**adaja
	Cʲ#C	ta**mʲ** **p**apɨ	ada**m** **pʲ**atava
	C#Cʲ	ta**m** **pʲ**apɨ	dina**mʲ** **p**adaja
	Cʲ#Cʲ	ta**mʲ** **pʲ**apɨ	dina**mʲ** **pʲ**atava
labial fricative-stop	C#C	ta**f** **p**apɨ	ʐira**f** **p**adaja
	Cʲ#C	ta**fʲ** **p**apɨ	potra**fʲ** **p**adaja
	C#Cʲ	ta**f** **pʲ**apɨ	ʐira**f** **pʲ**atava
	Cʲ#Cʲ	ta**fʲ** **pʲ**apɨ	potra**fʲ** **pʲ**atava
coronal stop-stop	C#C	ta**t** **t**apɨ	bra**t** **t**antsɨ
	Cʲ#C	ta**tʲ** **t**apɨ	bra**tʲ** **t**antsɨ
	C#Cʲ	ta**t** **tʲ**apɨ	bra**t** **tʲ**anuʃʲiva
	Cʲ#Cʲ	ta**tʲ** **tʲ**apɨ	bra**tʲ** **tʲ**anuʃʲiva
coronal nasal-stop	C#C	ta**n** **t**apɨ	kra**n** **t**antsa
	Cʲ#C	ta**nʲ** **t**apɨ	gra**nʲ** **t**antsa
	C#Cʲ	ta**n** **tʲ**apɨ	kra**n** **tʲ**anuʃʲiva
	Cʲ#Cʲ	ta**nʲ** **tʲ**apɨ	gra**nʲ** **tʲ**anuʃʲiva
coronal fricative-stop	C#C	ta**s** **t**apɨ	akra**s** **t**antsa
	Cʲ#C	ta**sʲ** **t**apɨ	ukra**sʲ** **t**antsɨ
	C#Cʲ	ta**s** **tʲ**apɨ	akra**s** **tʲ**anuʃʲiva
	Cʲ#Cʲ	ta**sʲ** **tʲ**apɨ	ukra**sʲ** **tʲ**anuʃʲix

4.2. Predictions

The *gestural blending hypothesis* predicts a blending of primary (Lips and TT) and secondary (TB) gestures. Particularly, we would expect the blending of Lips [critical, dental] of C1 and Lips [closed, labial] of C2 in labial fricative-initial sequences (/f$^{(j)}$#p$^{(j)}$/), as well as TT [critical, alveolar] of C1 and TT [closed, dental] of C2 in coronal fricative-initial sequences (/s$^{(j)}$#t$^{(j)}$/). Values for the Lips and TT in all the other sequences should not be affected, except for TT in C#Cʲ and Cʲ#C sequences (e.g. /n#tʲ/ and /nʲ#t/), where TT/TB blending is expected.

We would expect the blending of the conflicting TB gestures: the TB [narrow, velar] and TB [narrow, palatal] in either order for labial sequences (e.g. /m#pʲ/ and /mʲ#p/). A greater influence of TB [narrow, palatal] on the TB of the adjacent plain coronals might be expected, since these are assumed to be unspecified for TB (e.g. /s#tʲ/ and /sʲ#t/). The blending is expected to be manifested in a lower, more back, and relatively short TB gesture in Cʲ#C and C#Cʲ than in Cʲ#Cʲ (yet having a higher and more front position than in C#C). The timing of the TB gesture with the primary gesture should be different for all three sequences. Given the lower TB position of /sʲ/, we would expect that coronal fricative-initial sequences (e.g. /sʲ#t/ and /sʲ#tʲ/) would have lower (and possibly longer) TB than the other sequences.

Another prediction can be made about the relative TB magnitude and duration in C#Cʲ and Cʲ#C sequences based on the *gestural reduction hypothesis*. The syllable-final TB [palatal] gesture in Russian has been observed to reduce compared to the same gesture syllable-initially (Kochetov 2006a). We may, therefore, find differences in the magnitude and duration of TB in sequences C#Cʲ and Cʲ#C, regardless of gestural blending (e.g. /pʲ#p/ vs. /p#pʲ/).

4.3. Results and discussion

Mean TT (coronals) and TB (labials and coronals) magnitude values for stop-nasal-, and fricative-initial homorganic sequences are plotted in Figures 4 and 5. The values are labeled by the palatalization of the sequence (Cʲ#C, C#Cʲ, and C#Cʲ). The manner of articulation of C1 is not indicated, however, it should be noted that the most back values in Figure 4 and the lowest values for each sequence in Figure 5b represent coronal fricative-initial sequences (/s$^{(j)}$#t$^{(j)}$/). In Figure 5, TB values for target sequences can be compared to the TB position during the consonants /p/ (“control pp”) and /t/ (“control tt”) in control utterances. The results for the LA are not presented, since they were not significant. Relative timing of the TB gesture with respect to the primary gestures LA (a) and TT (b) in three different palatalization patterns, averaged across manner, are shown in Figure 6.

Contrary to the prediction of the *gestural blending hypothesis*, the results for the LA magnitude do not show the blending of Lips [critical, dental] and [closed, labial] in fricative-initial sequences (/f#p/, /f#pʲ/, /fʲ#pʲ/, and /fʲ#pʲ/). As before, fricatives are distinct from stops and nasals in their LA specification; none of the sequences are different in terms of LA from single consonants (in the context _#a, Experiment 1). The lack of blending is possibly due to a small overlap between the two gestures, as manifested in the substantially higher LA duration for sequences than for single consonants (in fact, the plateaus of sequences were almost 150% longer than the plateaus of single consonants).

Some support for the *gestural blending hypothesis* is provided by the results for the TT magnitude (Figure 4). The TT in sequences where one of the consonants is palatalized and the other is plain (Cʲ#C and C#Cʲ,e.g./nʲ#t/ and /n#tʲ/), is higher than in plain-plain sequences (C#C, e.g. /n#t/) and lower than in palatalized-palatalized sequences (Cʲ#Cʲ, e.g./nʲ#tʲ/). This is apparently due to the blending of TT [dental/alveolar] with TB [palatal], resulting in some retraction of the former gesture. Unexpectedly, however, fricative-initial sequences (e.g. /sʲ#t/ and /s#tʲ/) do not show such blending: as in the single condition, the TT during their production is higher and more back than in stops and nasals.

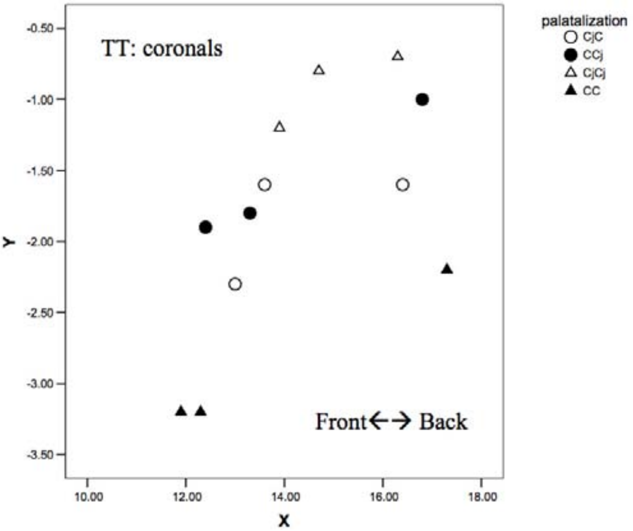
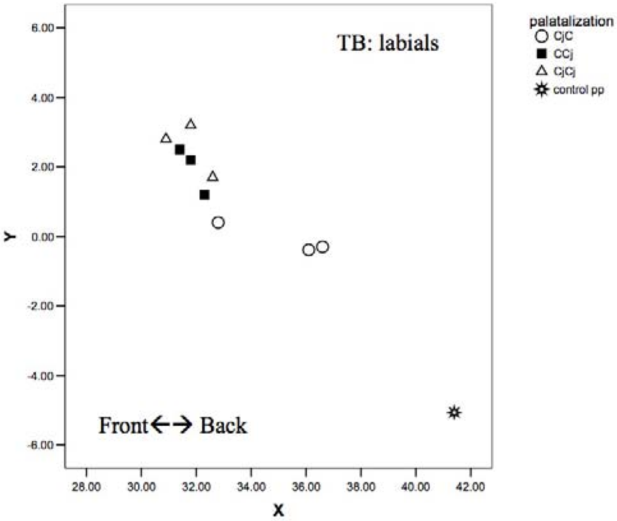
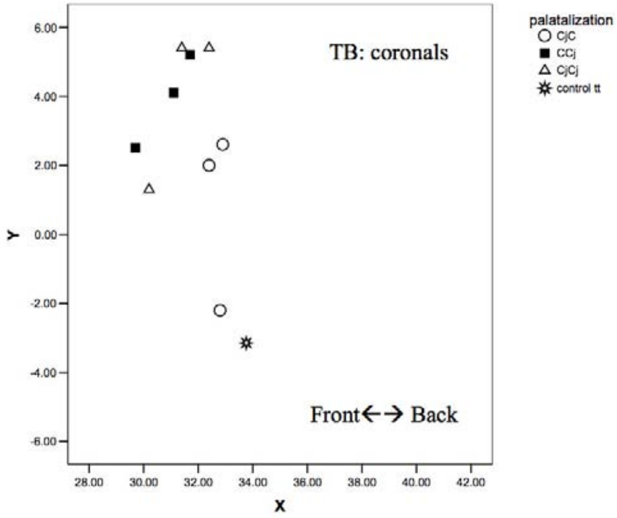


Figure 4: Mean Tongue Tip (TB) magnitude values (in mm) in homorganic coronal stop-, nasal-, and fricative-initial sequences that differ with respect to palatalization: Cʲ#C, C#Cʲ, Cʲ#Cʲ and C#C.

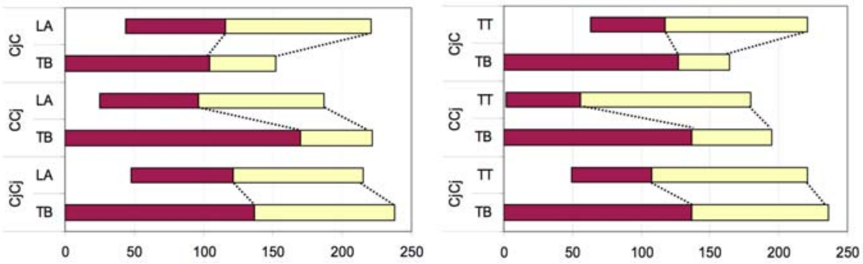


(a)



(b)

Figure 5: Mean Tongue Body (TB) magnitude values (in mm) in homorganic stop-, nasal-, and fricative-initial sequences that differ with respect to palatalization: C^j#C, C#C^j, and C^j#C^j, separately (a) for labials and (b) for coronals.



(a) (b)

Figure 6: Mean Lip Aperture (LA)/Tongue Tip (TT) and Tongue Body (TB) duration values in (ms) for labial (a) and coronal (b) homorganic stop-, nasal, and fricative-initial sequences that differ with respect to palatalization: C^j#C, C#C^j, and C^j#C^j.

As predicted by the *gestural blending hypothesis*, C^j#C sequences (e.g. /m^j#p/ and /s^j#t/) show a lower, more back, and shorter TB compared to the TB in C^j#C^j (e.g. /m^j#p^j/ and /s^j#t^j/), yet higher and more front than in C#C sequences (e.g. /m#p/ and /s#t/; Figures 5 and 6).^{4} This can be attributed to the blending of TB [narrow, velar] and TB [narrow, palatal]. However, the lack of significant difference in TB magnitude between C^j#C and C^j#a (e.g. /s^j#t/ vs. /s^j#a/ from Experiment 1; cf. Figure 2) suggests that blending may be responsible for the C^j#C vs. C^j#C^j differences in duration but not in magnitude. The latter difference can be attributed to the *gestural reduction*: lower/more back and shorter TB gesture is expected for syllable-final consonants in general (both in C^j#C and C^j#a). Interestingly, the expected blending in magnitude is not found for C#C^j sequences: the TB in this sequence is as high and front as in C^j#C^j (e.g. /m#p^j/ vs. /m^j#p^j/). In terms of duration, the TB of C#C^j is shorter than the TB of C^j#C^j, as one would expect (yet it is unexpectedly longer than the TB of C^j#C; Figure 6). This, again, could be in part attributed to gestural reduction. As expected, all three sequence types differed in the relative timing of TB with respect to the Lips and TT gestures. Further, the observed lower and longer TB in coronal fricative-initial sequences (Figure 6b) was expected (based on the results for single consonants).

5. Experiment 3

The goal of this experiment is to test the gestural reduction hypotheses by examining syllable-final coronal stops and nasals /t tʲ n nʲ/ before syllable-initial hetero-organic stops /p pʲ k kʲ/.

5.1. Method

Stimuli for this experiment are shown in Table 3. The control utterances shown at the bottom of the table were used to determine the position of the Tongue Tip in the absence of a coronal specification. Stimulus presentation was the same as in Experiments 1 and 2. A total of 180 tokens were collected (20 nonsense utterances * 5 repetitions + 20 real word utterances * 4 repetitions), or 9 tokens per each consonant in a given context.

The analysis included measurements of the Tongue Tip (TT) magnitude and duration, as defined in Experiment 1. For control items, TT magnitude was measured at ¼ of the acoustic closure of the sequence or at the midpoint of a single consonant. As before, duration measurements were limited to nonsense utterances. An ANOVA with factors C1 Manner (stop and nasal), C1 and C2 Palatalization (plain and palatalized), and C1 and C2 Place (labial and dorsal) was performed to determine significant effects and interactions. Another ANOVA with the factor Context was performed to compare the hetero-organic sequence context _#p with the single consonant condition (_#a).

5.2. Predictions

According to the *gestural reduction hypothesis* the TT [closed, dental] gesture is expected to be gradiently reduced (both in magnitude and duration) when followed by another oral gesture. It should, therefore, be different from the TT of the corresponding single consonants (_#a) (while both are expected to be reduced compared to syllable-initial gestures). The following dorsal consonants, /k/ (TB [closed, velar]) or /kʲ/ (TB [closed, velar]-TB [narrow, palatal]) might cause greater TT reduction, since the TT and TB are mechanically coupled. More specifically the reduction should be manifested in the backing and lowering of the TT as a consequence of TB backing. Further, the effect of this reduction might be smaller for palatalized coronals /tʲ nʲ/ for a similar reason: a coupling of the TT with the TB [narrow, palatal]. No differences in reduction are expected for stops and nasals (/t tʲ/ vs. /n nʲ/), since they have the same tract variable specifications.

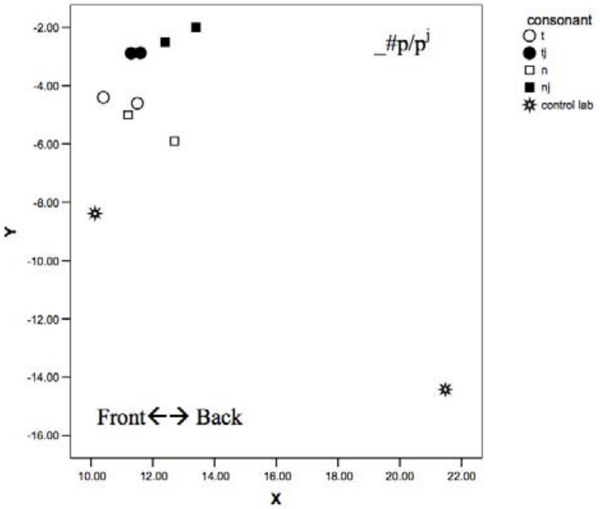
5.3. Results and discussion

Mean TT magnitude values for four coronal consonants are plotted in Figure 7 separately before labials (a) and before dorsals (b). The values can be compared to the TT position during the labial (“control lab”; the more back /p/ and the more front /pʲ/) and dorsal (“control dor”; the more back /k/ and the more front /kʲ/) in control nonsense and real stimuli. Plateau duration values are plotted in Figure 8.

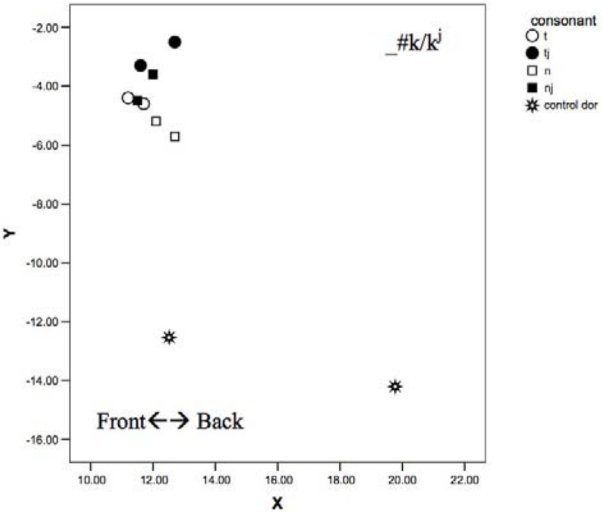
The TT magnitude results provide partial support for the *gestural reduction hypothesis*. As expected, the pre-consonantal TT reduction is greater before dorsals than before labials, however, this holds for nasals /n nʲ/ but not for stops /t tʲ/ (Figure 7). For the nasals, the reduction involves a relatively minor (about 1 mm) vertical displacement (TTy). Reduction in (plateau) duration is more extensive, and involves both stops and nasals (Figure 8). As expected, it is greater before dorsals /k kʲ/ was than before labials /p pʲ/. Unexpectedly, it is more extensive for palatalized consonants, /tʲ nʲ/, than for plain consonants, /t n/. Thus, the plateau of /nʲ/ before dorsals is almost twice as short as before labials (See Figure 8). A comparison with the single consonant condition (_#p vs. _#a from Experiment 1), shows a very small (yet significant) reduction in magnitude: the TT is lower than before /a/ by 0.5 mm. The pre-consonantal context is also shorter in duration, however, this applies to stops but not to nasals.

Table 4: Stimuli used in Experiment 3.

Consonant	Context	Nonsense word stimuli	Real word stimuli
/t/	__#p	ta**t** **p**apɨ	bra**t** **p**adaja
	__#pʲ	ta**t** **pʲ**apɨ	bra**t** **pʲ**atava
	__#k	ta**t** **k**opɨ	bra**t** **k**adatʃnʲika
	__#kʲ	ta**t** **kʲ**opɨ	bra**t** **kʲ**edra
/tʲ/	__#p	ta**tʲ** **p**apɨ	bra**t** **p**adaja
	__#pʲ	ta**tʲ** **pʲ**apɨ	bra**t** **pʲ**atava
	__#k	ta**tʲ** **k**opɨ	bra**t** **k**adatʃnʲika
	__#kʲ	ta**tʲ** **kʲ**opɨ	bra**t** **kʲ**edra
/n/	__#p	ta**n** **p**apɨ	kra**n** **p**adaja
	__#pʲ	ta**n** **pʲ**apɨ	kra**n** **pʲ**atava
	__#k	ta**n** **k**opɨ	kra**n** **k**adatʃnʲika
	__#kʲ	ta**n** **kʲ**opɨ	kra**n** **kʲ**edra
/nʲ/	__#p	ta**nʲ** **p**apɨ	gra**nʲ** **p**adaj
	__#pʲ	ta**nʲ** **pʲ**apɨ	gra**nʲ** **pʲ**atava
	__#k	ta**nʲ** **k**opɨ	kra**nʲ** **k**adatʃnʲika
	__#kʲ	ta**nʲ** **kʲ**opɨ	gra**nʲ** **kʲ**edra
control /p/	__#p	ta**p** **p**apɨ	kra**p** **p**adaja
control /pʲ/	__#pʲ	ta**pʲ** **pʲ**apɨ	gra**pʲ** **pʲ**atava
control /k/	__#k	**k**opɨ	**k**adatʃʲnʲik
control /kʲ/	__#kʲ	**kʲ**opɨ	**kʲ**edrɨ



(a)



(b)

Figure 7: Mean Tongue Tip (TT) magnitude values (in mm) of coronal stops and nasals /t t^j n n^j/ before hetero-organic consonants: (a) /p p^j/ and (b) /k k^j/.

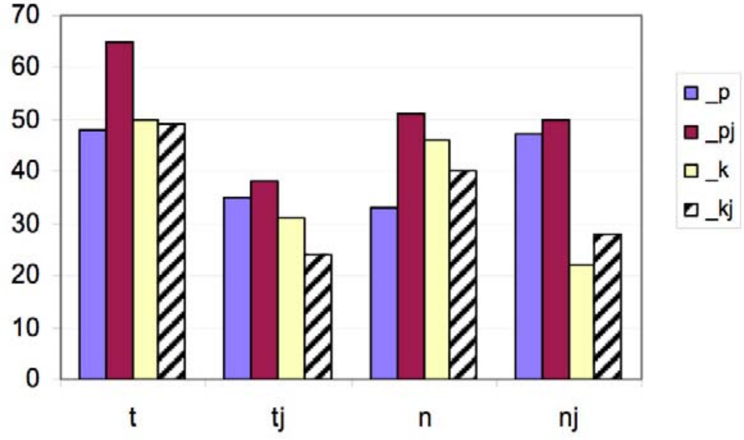


Figure 8: Mean Tongue Tip (TT) plateau duration values (in ms) for coronal stops and nasals /t t^j n n^j/ before hetero-organic consonants /p p^j k k^j/.

The observed degree of magnitude reduction is rather small compared to the extensive TT reduction often reported for English (Browman and Goldstein 1995; Byrd 1996a, among others). Also surprising is the robust difference between the rather small degree of magnitude reduction and a rather large degree of duration reduction (or truncation; cf. Kochetov 2006a on the Lip duration reduction).

6. General discussion and conclusions

Table 5 summarizes gestural specifications for the consonants, modified based on our results (cf. Table 1), adding values for stiffness, *k* (as a predictor of duration). Constriction Degree (CD), Constriction Location (CL), stiffness, and phasing relations are indexed by numbers, indicating whether variables are the same or different for different consonants. Values that were not possible to determine are omitted ("…"). Differences between consonants in values that were not predicted given our assumptions in Section 2.1 are shown in bold.

The results show that the assumed gestural representations (see Table 1) capture much of the variation observed in the production of Russian consonants. Gestures such as, for example, the Lips and TT, are found to be largely invariant in terms of their constriction degree (CD) and constriction location (CL), with values being similar for different consonants (e.g. stops and nasals). In Table 5, this is indicated by the same CD and/or CL values for

Table 5: Specifications of articulatory gestures of Russian plain/palatalized labial and coronal stops, nasals, and fricatives suggested by the results of the study.

Cons.	Primary constriction gesture	Secondary constriction gesture	Phasing
/p/	Lips [closed$_1$, labial$_1$, **k$_1$**]	TB [**narrow$_1$**, velar$_1$, ...]	...
/m/	Lips [closed$_1$, labial$_1$, **k$_2$**]	TB [**narrow$_2$**, velar$_1$, ...]	...
/p^j/	Lips [closed$_1$, labial$_1$, **k$_1$**]	TB [**narrow$_3$**, palatal$_1$, **k$_1$**]	(TB-Lips)$_1$
/m^j/	Lips [closed$_1$, labial$_1$, **k$_2$**]	TB [**narrow$_4$**, palatal$_1$, **k$_1$**]	(TB-Lips)$_1$
/f/	Lips [critical$_1$, dental$_1$, k$_3$]	TB [narrow$_2$, velar$_1$, ...]	...
/f^j/	Lips [critical$_1$, dental$_1$, k$_3$]	TB [narrow$_4$, palatal$_1$, **k$_2$**]	(TB-Lips)$_1$
/t/	TT [closed$_2$, dental$_2$, **k$_4$**]	*n/a*	*n/a*
/n/	TT [closed$_2$, dental$_2$, **k$_5$**]	*n/a*	
/t^j/	TT [closed$_2$, dental$_2$(?), **k$_6$**]	TB [**narrow$_5$**, palatal$_2$, **k$_3$**]	**(TB-TT)$_1$**
/n^j/	TT [closed$_2$, dental$_2$(?), **k$_7$**]	TB [**narrow$_5$**, palatal$_2$, **k$_3$**]	**(TB-TT)$_1$**
/s/	TT [critical$_2$, alveolar$_1$, k$_8$]	*n/a*	*n/a*
/s^j/	TT [critical$_2$, alveolar$_1$, k$_8$]	TB [**narrow$_6$**, palatal$_2$, **k$_4$**]	**(TB-TT)$_2$**

classes of consonants: the labial stops and nasals /p m p^j m^j/ have the same specification for Lips CD and CL ([closed, labial]), and different from the labial fricatives /f f^j/; similarly, coronal stops and nasals /t n t^j n^j/ are specified identically for TT CD and CL ([closed, dental]), and differently from the coronals fricatives /s s^j/. (The CL [dental] specification for /t^j n^j/ assumes that its blending with TB can produce the actual alveolar realization. An alternative TT specification for these consonants is [closed$_2$, alveolar$_2$].) These similarities and differences are fully predicted by the model (see Section 2.1).

Other gestures, such as the TB, vary considerably in their CD and CL depending on which other gesture they are coordinated with. Note that in Table 5, the TB CD specification of plain and palatalized labial stops is different from the corresponding nasals (and fricatives). Among the coronals, the difference in TB CD is among stops/nasals and fricatives. These differences are not fully expected (cf. Table 1). It is possible that the key difference here is between consonantal and vocalic gestures (Browman and Goldstein 1992: 164–165), with secondary articulations falling in the latter class.

Further, all the examined gestures show certain differences in duration that are not captured by the assumed specifications. In Table 5, these differences are assumed to result from different stiffness (k) specifications. Note that labial stops, nasals, and fricatives have distinct Lips stiffness specifications (yet the same for plain and palatalized counterparts). Among the coronals, TT stiffness specifications differ by manner and secondary articulation (except for fricatives). For palatalized consonants, the TB stiffness is different for labials and coronals, and within these classes for stops/nasals. This variability appears to correspond to higher-level segmental categories (e.g. stops vs. nasals). Thus, although gestures as ‘atomic’ units combine to form segments, the derived higher-level categories appear to constrain the variation at the lower level – a common property of self-organizing systems (Kauffman 1995). The questions remain, however, why segmental information affects primarily duration (stiffness) and how such influences can be captured in the current Gestural Linguistic Model.

As expected, the phasing specification is the same for all palatalized labials, syllable-final /pʲ mʲ fʲ/: the secondary constriction is phased slightly prior to the primary constriction (Kochetov 2006a). The phasing specification of the palatalized coronals (syllable-final /tʲ nʲ sʲ/), however, is different and sensitive to the manner of articulation of the primary gesture. While for /tʲ nʲ/, the secondary constriction is phased at the release of the primary constriction, the two are phased simultaneously for /sʲ/ (with the TB constriction being formed prior to and released after the TT constriction). The difference between coronals and labials in phasing can be attributed to the presence or absence of coupling of the two articulators. The difference within coronals, however, may follow from the considerable durational differences between fricatives and stops/nasals. This difference may also have an acoustic explanation: the lack of the TB movement during the /sʲ/ primary constriction is necessary to maintain the stable high frequency noise pattern characteristic of this fricative.

The results of the study also raise a question about how to explain variation in degrees of gestural blending or unexpected absence of it. Thus, the within-segment TT/TB blending in /tʲ nʲ/ affects CL but not CD, with CL relatively slightly affected, producing laminal alveolars rather than post-alveolars. The TT of /sʲ/ is resistant to any blending (possibly due to the contrast with two post-alveolar fricatives /ʂ/ and /ʃ/). In consonant sequences, the Lips are strongly resistant to blending, while the TT and TB are somewhat affected. To capture this variability in blending, it seems to be necessary to specify relative *weights* for tract variables. More research is needed to determine to what extent such weights are inherent, language-independent properties of specific tract variables and to what extent they are specified on a language-particular basis.

Finally, the results of the study provide evidence for gestural reduction, while showing that its manifestations are somewhat different for different gestures (and possibly segments). Thus, the Tongue Body shows considerable reduction, while the Tongue Tip reduces slightly in magnitude and considerably in duration (cf. Kochetov 2006a on TB and Lips). This, again, underscores the apparently rather loose relation between the magnitude and duration of gestures, suggesting that stiffness may not be the only factor that affects gestural duration. Differences in the degree of magnitude reduction will possibly require different damping ratio specifications.

To conclude, the current paper tested a number of specific predictions made by the Gestural Linguistic Model of Articulatory Phonology by examining production of Russian consonants in several contexts. The results provided support for the Gestural Linguistic Model in general, showing a relatively good match between predicted and observed variation in articulatory movements. At the same time, the results highlighted certain weaknesses of the model, which under-predicted some specific kinds of articulatory variability (variability in duration and reduction patterns) and over-predicted others (variability in terms of gestural blending). The results also revealed some interesting and partly unexpected interactions between articulatory phonetic units (gestures) and phonological categories (segments and the syllable). Taken together, the findings of the current study suggest that certain modifications are necessary so that the Gestural Linguistic Model is capable of generating a fuller range of attested patterns of gestural variation. First, there is a need for a more realistic modeling of variation in gestural duration and its relation with other gestural properties (cf. Byrd 1996b). Second, the model seems to require more detailed gestural specifications with particular tract variables assigned relative weights, and possibly incorporating some notion of *degrees of articulatory constraint* (DAC: Recasens and Pallarès 2001). Third, the traditional assumption of AP that segments as phonological units are purely epiphenomenal (given that gestures are units of contrast) may need to be revised. It appears that both gestures and segments (as well as prosodic structure) have to be part of a model of speech production. A revised model, therefore, would need to allow for a principled interaction of higher-level segmental and lower-level gestural units. Further development of language-particular versions of the Gestural Linguistic Model will help researchers address these issues in greater depth and will therefore provide better insight in the sources of phonetic variation.

Acknowledgements

The author would like to acknowledge Louis Goldstein's considerable contribution to this study. Thanks to an anonymous reviewer for helpful comments and insightful suggestions. Thanks also to the organizers and the audience of the conference Variation and Change in Phonology and Phonetics, Potsdam, Germany, October 2004. All errors are my own. The research was supported by Postdoctoral Fellowship # 756-01-0145 from the Social Sciences and Humanities Research Council of Canada (SSHRC) held at Haskins Laboratories, New Haven, CT. All errors are my own.

Notes

1. Some of the stimuli used in the study (10 items with /p/ and /p^j/) have been previously employed in an articulatory study of syllable position effects (Kochetov 2006a).
2. Consistent duration measurements of the TB and TT in real words and TB [velar] (plain labials) in all items were not possible due to frequent interference from preceding or following consonants (especially /r/).
3. Somewhat unexpectedly, LA values were smaller for fricatives than for stops, due to the greater raising of the LL (that apparently accompanied its movement to the upper teeth) (see Figure 2a).
4. TB magnitude values for coronal fricative-initial sequences were close to the control (the lowest C^{j}C value in Figure 5b).

References

Bolla, K.
1981 *A Conspectus of Russian Speech Sounds*. Cologne: Bölau.

Browman, Catherine P. and Louis Goldstein
1989 Articulatory gestures as phonological units. *Phonology* 6: 201–252.

Browman, Catherine P. and Louis Goldstein
1990 Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics* 18: 299–320.

Browman, Catherine P. and Louis Goldstein
1992 Articulatory phonology: An overview. *Phonetica* 49: 155–180.

Browman, Catherine P. and Louis Goldstein
1995 Gestural syllable position effects in American English. In *Producing Speech: Contemporary Issue*s *(for K. S. Harris),* F. Bell-Berti and L. J. Raphael (eds.), 19–33. Woodbury, N.Y.: AIP Press.

Byrd, Dani
1996a Influences on articulatory timing in consonant sequences. *Journal of Phonetics* 24(2): 209–244.

Byrd, Dani
1996b A phase window framework for articulatory timing. *Phonology* 13(2): 139–169.

Byrd, Dani
2003 Frontiers and challenges in Articulatory Phonology. *International Congress of Phonetic Sciences* (Barcelona, Spain): 1–4.

Gick, Bryan
2003 Articulatory correlates of ambisyllabicity in English glides and liquids. In *Papers in Laboratory Phonology VI: Constraints on Phonetic Interpretation,* J. Local, R. Ogden, and R. Temple (eds.), 222–236. Cambridge: Cambridge University Press.

Goldstein, Louis, Dani Byrd, and Elliott Saltzman
2006 The role of vocal tract gestural action units in understanding the evolution of phonology. In *From Action to Language: The Mirror Neuron System,* M. Arbib (ed.), 215–249. Cambridge: Cambridge University Press.

Kauffman, Stuart
1995 *At Home in the Universe: The Search for the Laws of Self-organization and Complexity*. Oxford: Oxford University Press.

Keating, Patricia A.
1990 The window model of coarticulation: Articulatory evidence. In *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech,* J. Kingston and M. E. Beckman (eds.), 451–470. Cambridge: Cambridge University Press.

Kochetov, Alexei
2006a Syllable position effects and gestural organization: Articulatory evidence from Russian. In *Papers in Laboratory Phonology VIII,* Louis Goldstein, Douglas Whalen, and Catherine Best (eds.), 565–588. Berlin/New York: Mouton de Gruyter.

Kochetov, Alexei
2006b Testing Licensing by Cue: A case of Russian palatalized coronals. *Phonetica* 63: 113–148.

Krakow, Rena A.
1999 Physiological organization of syllables: A review. *Journal of Phonetics* 27: 23–54.

Labov, William
2001 *Principles of Linguistic Change: Social Factors*. Oxford: Blackwell.

Lindblom, B.
1990 Explaining phonetic variation: A sketch of the H-H theory. In *Speech*

Production and Speech Modeling, William Hardcastle and Alain Marchal (eds.), 403–439. Kluwer Academic Publishers, Dordrecht.

Perkell, J. S., M. Cohen, M. Svirsky, M. Matthies, I. Garabieta, and M. Jackson
1992 Electro-magnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. *Journal of the Acoustical Society of America* 92: 3078–3096.

Recasens, D. and M. D. Pallarès
2001 V-to-V coarticulation in Catalan VCV sequences. *Journal of the Acoustical Society of America* 76: 1624–1635.

Saltzman, Elliott and Dani Byrd
2000 Task-dynamics of gestural timing: Phase windows and mulitfrequency rhythms. *Human Movement Science*, 19: 499–526.

Zsiga, Elizabeth C.
1995 An acoustic and electropalatographic study of lexical and postlexical palatalization in American English. In *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence*, B. Connell and A. Arvaniti (eds.), 282–302. Cambridge: Cambridge University Press.

Variation in the perception of an L2 contrast: A combined phonetic and phonological account[1]

Silke Hamann

The present study argues that variation across listeners in the perception of a non-native contrast is due to two factors: the listener-specific weighting of auditory dimensions and the listener-specific construction of new segmental representations. The interaction of both factors is shown to take place in the perception grammar, which can be modelled within an OT framework. These points are illustrated with the acquisition of the Dutch three-member labiodental contrast [ʋ v f] by German learners of Dutch, focussing on four types of learners from the perception study by Hamann and Sennema (2005a).

1. Introduction

Dutch has a cross-linguistically very unusual labiodental contrast between an approximant /ʋ/, a voiced fricative /v/, and a voiceless fricative /f/ (see e.g. Booij 1995 and Gussenhoven 1999). Minimal triplets illustrating the contrast are given in (1).

(1)

wee	vee	fee	[ʋeː, veː, feː]	‘contraction, cattle, fairy’
wijj	vijl	feil	[ʋɛil, vɛil, fɛil]	‘while, rasp, error’

The occurrence of this contrast is restricted to word-initial position between voiced segments. Many speakers of Standard Dutch, especially those from the western part of the Netherlands, neutralize the /f/ - /v/ contrast (Booij 1995: 7f.). According to Mees and Collins (1982: 5), no neutralization occurs in careful speech and in more formal registers. Very Southern Dutch and Flemish speakers realise the approximant as bilabial (Booij 1995: 8), which can be transcribed as /β̞/ (see Gussenhoven 1999: 75).[2] This regional variance is not further considered in the following study.

German learners of Dutch have problems acquiring the three-member contrast since their native language differentiates only a voiced and a voiceless labiodental fricative (Kohler 1999; Wiese 1996), see the minimal pairs in (2).[3]

(2)

Wein	fein	[vain, fain]	‘wine, fine’
Wort	fort	[vɔɐt, fɔɐt]	‘word, away’

In the phonetic literature on German, the voiced labiodental sound is usually described as a fricative (see e.g. Jessen 1998, Kohler 1999, Wängler 1974). However, Kohler (1995: 154) mentions that German /v/ can turn into an approximant, especially in initial position, and Viëtor (1897: 229, 231) writes that friction in German /v/ is very little compared to English and French /v/.

Several studies on second language (L2) acquisition have shown that the existence of the same phonological category in the first language (L1) does not necessarily result in perfect L2 performance (see e.g. Best and Strange 1992; Flege 2002; Iverson et al. 2003). Instead, the phonetic realisations of the respective categories play a major role in determining the categorization. Hamann and Sennema (2005a) add to this evidence with a categorisation experiment that tested the perception of the Dutch labiodental contrast by German learners of Dutch. They found that the German learners had problems perceiving the Dutch /v/ correctly, despite the fact that German has the same phonological category.

The present article gives an account of the data acquired in Hamann and Sennema's study, describing the performance of four types of speakers. The data is used to illustrate that L2 perception has to be accounted for with the fine-grained auditory differences between L1 and L2 segments *and* the influence of the native phonology (see e.g. Flege and Hillenbrand 1986). This interaction of phonetics and phonology can be modelled with language-specific perception grammars, which map auditory forms onto surface phonological forms. Proposals for such perception grammars have been made by Boersma (1998 et seq.) and Pater (2004) in the framework of Optimality Theory (McCarthy and Prince 1993, Prince and Smolensky 1993; henceforth: OT). In the present article, the perception grammars of the L1 and L2 native speakers and of the L2 learners are formalised. Furthermore, the article illustrates that the differences between the native and the L2 perception grammar allow different acquisition strategies, mirrored in the varied perceptual performance of

L2 learners. Thus, variation in the perception of an L2 contrast is accounted for by differences in the L2 perception grammars.

The article is structured as follows. Section 2 presents data on the performances by four groups of Dutch learners from Hamann and Sennema's perception study. In section 3, acoustic differences between the labiodentals in Dutch and German are discussed. Section 4 formalises the perception grammar of Dutch and German native listeners, and section 5 the perception grammars of the four types of learners of Dutch, illustrating different acquisition strategies. Section 6 summarises and concludes.

2. Variation in the perception of a second language: four types of learners

In a categorisation experiment, Hamann and Sennema (2005a) tested the perception of the Dutch labiodental contrast by three groups: six German listeners without any knowledge of Dutch (German L1), twenty-one German learners of Dutch (Dutch L2), and six Dutch native listeners (Dutch L1). The test materials did not only contain the three labiodentals but all Dutch obstruents /p, b, t, d, k, x, f, v, ʋ, s, z, ɕ/ (see Mees and Collins 1982).[4] These consonants were followed by the vowel /a/ and were read eight times by a male speaker in the sentence "Hoor je __", 'Do you hear __'. The token sentences were repeated four times and the total set of 384 stimuli sentences were randomized. The participants were presented with one stimuli sentence at a time via headphones and had to click on orthographic representations of the consonant-vowel sequences. With respect to the labiodentals, the German L1 listeners had to classify the three Dutch labiodental sounds as one of their two native German sounds /f/ and /v/ or any other of the ten obstruents /p, b, t, d, k, g, s, z, ç, ʃ/ in German (Kohler 1999). The Dutch native and L2 listeners had as answer categories the 12 Dutch consonants /p, b, t, d, k, x, f, v, ʋ, s, z, ɕ/. The categorisation results for the labiodentals, split by the three participant groups, are given in table 1 on the next page.

Based on the phonological descriptions of the categories alone, we would expect the German listeners to equate the Dutch /f/ with their /f/, the Dutch /v/ with their /v/, thus to perform well for these two known categories, and to show problems with Dutch /ʋ/, since German does not have an equivalent category for this sound. The results of the experiment in table 1 illustrate

Table 1: Mean identification scores (percent correct) of the three test groups for the labiodentals in the perception experiment by Hamann and Sennema (2005a), with stimuli in rows, and responses sorted by language group in columns. The numbers in each row per language group do not add up to 100 percent, because miscategorisations involving non-labiodental sounds are not included.

		German L1		Dutch L2			Dutch L1		
		/f/	/v/	/f/	/v/	/ʋ/	/f/	/v/	/ʋ/
stimulus	/f/	99.5	0	79.0	17.7	2.1	94.8	5.2	0
	/v/	16.7	82.8	5.2	74.6	18.5	5.2	94.8	0
	/ʋ/	0	99.5	0.1	6.1	92.6	0	0	99.5

that the categorization of the Dutch sounds by German listeners departs from these expectations.

German native listeners, without knowledge of Dutch, perceived the Dutch voiced fricative as their voiced fricative in 82.8 percent of the cases, and the Dutch labiodental approximant /ʋ/ as their voiced fricative /v/ in almost all of the cases, converse to the phoneme-based prediction. Unsurprisingly, Dutch /f/ was perceived as German /f/ in 99.5 percent of the cases. The *German learners of Dutch* were also not in line with the expectations. They succeeded in categorising Dutch /ʋ/ correctly, but miscategorised /f/ in 21 percent of the cases and /v/ in 25.4 percent of the cases.

The performance of the listeners in this study varied depending on the speakers and on their level of proficiency. Whereas certain mis-categorisations almost never occurred (e.g. the categorisation of /f/ as /ʋ/ and reverse), others were made by a large number of participants (e.g. confusing /v/ and /ʋ/). A detailed examination of the data showed that recurring patterns in the performance of the participants allow us to group them into four types of learners (miscategorisations below 10 percent are not taken into consideration).[5]

The first type of learners, termed Learner A in the following, does not discriminate Dutch /v/ and /ʋ/, but categorises tokens of both types of sound as approximant, see table 2. Though this pattern could only be observed for two of the twenty-one participants (PP5 and PP11)[6], we will see below that it is of great interest for modelling the learning process because it shows the application of a German perception grammar to the Dutch sounds, i.e. the initial stage of L2 learning. Three participants (PP1, PP20, PP21) are summarized here as learner type B because they show almost identical patterns of confusing /f/ and /v/.[7] A similar mis-categorisation of /f/ as /v/ could also be observed for seven other participants (PP7, PP11, PP12, PP14, PP15, PP17 and PP18). A third group of participants (PP3, PP6, PP9) confused /v/ and /ʋ/. This type of learning is summarized as Learner C. Again, we find part of their miscategorisation patterns in other participants: the misperception of /v/ as /ʋ/ by nine participants (PP3, PP5-PP9, PP13, PP15 and PP17), and the misperception of /ʋ/ as /v/ by two participants (PP8 and PP14). Five participants (PP2, PP4, PP10, PP16 and PP19) performed native-like in making almost no mistakes in the categorisation of all three labiodentals. These participants are represented by learner type D. A correct identification of /f/ could be observed for further five participants (PP3, PP5, PP6, PP9 and PP13), a correct identification of /v/ for three (PP12, PP14 and PP18), and a correct identification of /ʋ/ for ten participants (PP5, PP7, PP11-PP13, PP15, PP17, PP18, PP20 and PP21).

These four types of learners represent 13 of the 21 participants. The remaining eight listeners are covered by combinations of two learning types, with one exception: participant PP8 categorised /f/ as /ʋ/ (in 40.6 percent of the cases), a miscategorisation that none of the other listeners showed and that is not further considered in the present study.

Table 2: Identification scores (percent correct) for four different types of learners in the study by Hamann and Sennema (2005a: 166).

		Learner A			Learner B		
		/f/	/v/	/ʋ/	/f/	/v/	/ʋ/
stimulus	/f/	100%	0%	0%	40%	40%	0%
	/v/	0%	0%	100%	60%	60%	0%
	/ʋ/	0%	0%	100%	0%	0%	100%

		Learner C			Learner D		
		/f/	/v/	/ʋ/	/f/	/v/	/ʋ/
stimulus	/f/	100%	0%	0%	100%	0%	0%
	/v/	0%	75%	25%	0%	100%	0%
	/ʋ/	0%	25%	75%	0%	0%	100%

The performance of these four types of learners cannot be accounted for by a comparison of the phonemic categories in both languages, because in that case we would expect the German listeners to have problems with the Dutch labiodental approximant, a category that does not exist in German. However, only Learner C showed this pattern. Conversely, three of the four learner types had problems categorising the Dutch voiced labiodental fricative /v/ correctly, although this category exists in German. An alternative explanation for the performance of the German listeners is the (dis)similarity in the *phonetic realisations* of these categories, which is the topic of the following section.

3. Acoustic and auditory differences between German and Dutch labiodentals

The labiodentals in Dutch and German contrast in voicing (voiced versus voiceless fricatives), and the Dutch ones additionally in manner (fricative versus approximant). Acoustically, fricatives differ from approximants in their presence of friction noise, absence of continuous formants and a longer duration. Voiced fricatives can be distinguished from voiceless ones by the presence of a voicing bar and periodicity of the signal. Usually, voiced fricatives are also shorter than their voiceless counterparts, see Stevens et al. (1992) for English, Mees and Collins (1982) for Dutch and Jessen (1998) for German. *Duration* is thus employed both for distinguishing fricatives from approximants and voiced from voiceless fricatives. Furthermore, duration is a well-known auditory dimension. Humans are able to perceive small durational differences in the speech signal due to the high temporal resolution of their auditory system (Plack 2004: 19).

The presence or absence of a voicing bar and of continuous formants, on the other hand, are acoustic characteristics but not necessarily relevant auditory dimensions. Instead, listeners seem to pay attention to the periodicity versus aperiodicity of a signal (Faulkner and Rosen 1999). The presence of friction noise is closely related to the auditory dimension of periodicity, since a segment with a large friction component is less periodic. An acoustic measure for this relation is the harmonics-to-noise ratio, or *harmonicity median*, which indicates the ratio of periodicity to friction in a sound. A signal with a harmonicity median of 0 dB, for instance, has equal energy in the harmonics and in the noise, and a signal with a harmonicity median of 20 dB has almost 100% of the energy in the periodic part (Boersma 1993). In the following, we employ harmonicity median as auditory dimension of periodicity and friction.

In an acoustic study with five Dutch and five German female speakers, Hamann and Sennema (2005b) compared, among other parameters, the duration and harmonicity median of the German labiodentals /f, v/ with those of the Dutch labiodentals /f, v, ʋ/. The results of this study show that both German /v/ and Dutch /ʋ/ share a short duration and a high harmonicity median (i.e., they have little friction and are voiced). Dutch /v/ differs in both parameters; it is of medium duration and has a low harmonicity median (i.e., has considerable friction and is voiced). Dutch and German /f/ are both of long duration and have a negative harmonicity median (indicating considerable friction and voicelessness). These results are presented in figure 1.

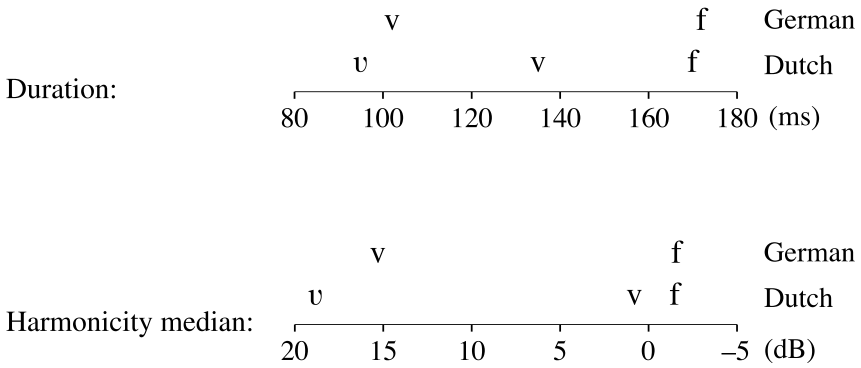


Figure 1: Scales comparing the realisations of the two German and the three Dutch labiodentals in the acoustic study by Hamann and Sennema (2005b) along the dimensions of duration (top) and harmonicity median (bottom).

The German labiodental fricative /v/ is acoustically closer to the Dutch labiodental approximant /ʋ/ than to the corresponding Dutch labiodental fricative /v/. These findings are in line with the observations made by Viëtor (1897) and Kohler (1995) that German /v/ is not very fricative-like and can be realised as an approximant in initial position.

The actual spreading of the labiodental tokens from Hamann and Sennema's data along duration and harmonicity median are shown in figure 2, where we can see a clear overlap of Dutch /ʋ/ and German /v/, and of Dutch and German /f/. The large variation in the realisation of Dutch /v/ and its partial overlap with Dutch /f/ is probably due to the often occurring

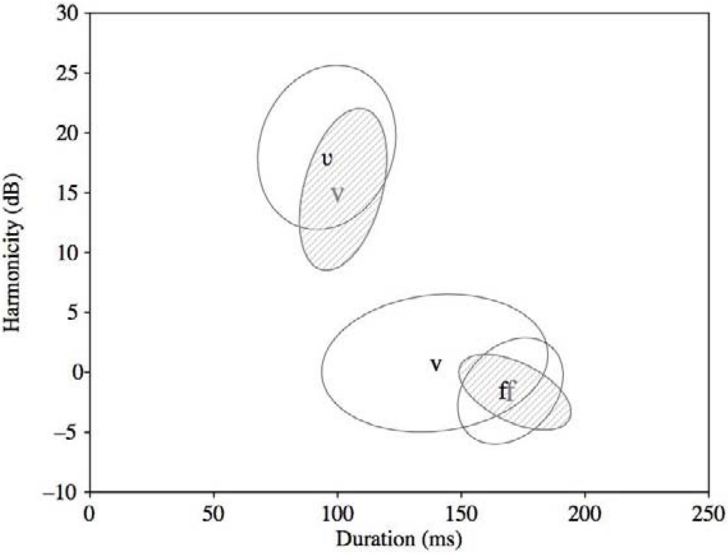


Figure 2: Distributions of German labiodentals (grey lines and hatched) and Dutch labiodentals (black lines) and their average values (position of symbols) along the dimensions of duration in ms (horizontal) and harmonicity median in dB (vertical).

neutralisation of /v/ to /f/ described in section 1, though the speakers in the study were from the area of Nijmegen, where a contrast is supposedly made.

In contrast to the average values for the five Dutch speakers in figure 2, the stimuli produced by the Dutch male speaker (also from the area of Nijmegen) that were used in the perception experiment, given in figure 3, did not show a three-way distinction in duration for the labiodentals. Instead, the realisations of /v/, in the lower middle of figure 3, and those of /ʋ/, in the upper part of figure 3, are almost of equal duration for this speaker.

The comparison of the average distributions for both languages (figure 2) offers an explanation for the perceptual misclassifications of Dutch labiodentals by the German L2 learners described in section 2. As we can infer from the grey ellipsis on the upper left in figure 2, German listeners are used to encounter tokens of /v/ with a duration between 85 and 120 ms and a harmonicity median between 9 and 22 dB. The realisations of Dutch /ʋ/ are very similar to those. Likewise, the realisations of Germans /f/ with a

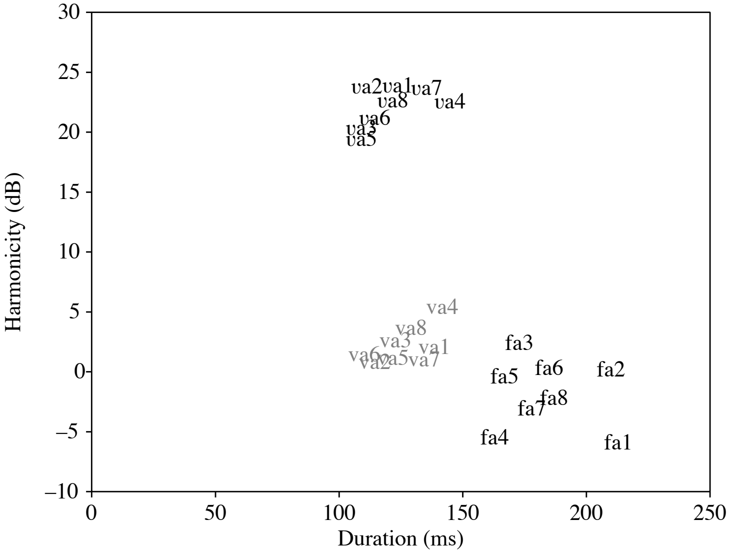


Figure 3: Dutch labiodental tokens (by one speaker) used in the perception experiment along the dimensions of duration (ms) and harmonicity (dB): tokens of /ʋ/ in the upper middle, tokens of /v/ in the lower middle, and those of /f/ in the lower right corner.

duration between 150 and 200 ms and a harmonicity median between –5 and 2 dB are very similar to those of Dutch /f/. Thus Germans can employ the same auditory dimensions and very similar values along these dimensions for categorising Dutch /ʋ/ and /f/ than they use for categorising their native labiodentals. However, tokens of Dutch /v/ pose a problem for German listeners because, although these tokens partly overlap with the two categories existing in German, they also provide values in between that the German listeners are not acquainted with.[8] More importantly, the German listeners do not have a third category to associate these values with. This accounts for the problems German listeners have with correctly categorizing the Dutch voiced fricative. German learners of Dutch have to create a new segmental category for the Dutch /v/ and to associate the relevant auditory values with it, before they are able to classify the Dutch labiodentals correctly. The following section shows how the association of (new) segmental categories with auditory dimensions and specific values along these dimensions can be formalised.

4. Perception grammars for German and Dutch listeners

In speech perception, an auditory input is mapped onto an abstract phonological surface form. This process is largely language-specific, as both the abstract categories and their concrete auditory realisations can differ from language to language. The involvement of language-specific, i.e. grammatical, knowledge in speech perception has lead a number of researchers to model this process with a *perception grammar* within an OT framework, see Boersma (1998 et seq.) and Pater (2004) for L1, and Hayes (2002), Escudero and Boersma (2003, 2004) and Escudero (2005) for L2.

The mapping of auditory forms onto phonological forms can be formalised in an OT perception grammar by means of negative *cue constraints*, as proposed by Boersma (1998) and Escudero and Boersma (2003):

(3) *Cue constraints (or perceptual construction constraints)*:

> "A value *x* on the auditory continuum f should not be perceived as the phonological category *y*" or short "*xf* not /*y*/".

In our case, we use for instance the cue constraint "110 ms not /f/" to formalise the fact that a segment of 110 ms duration is not a voiceless labiodental fricative (at least at normal speaking rate)[9], neither in German nor in Dutch.

Every cue constraint that maps an auditory value onto a phonological category has antagonistic cue constraints that map the same value onto other phonological categories in the respective language. Staying with our example and restricting our description to labiodentals, we also have, e.g., "110 ms not /v/" for German. Since 110 ms is a perfectly acceptable duration for a /v/ (but not for a /f/) in German, the constraint "110 ms not /v/" is much lower ranked than its antagonist "110 ms not /f/", and therefore plays only a small role in the decision on the perceived category (due to the nature of OT as a decision-by-exclusion mechanism). Dutch has the additional constraint "110 ms not /ʋ/". From the data in figure 2 we can see that a /v/ with a duration of 100 ms does occur in Dutch, but is a marginal case of this category. For this reason, "110 ms not /v/" is higher ranked than "110 ms not /ʋ/", but lower than "110 ms not /f/", to mirror the fact that a token with 110 ms is usually classified as /ʋ/, sometimes as /v/, but never as /f/. The actual decision on the classification hinges then on the evaluation noise and the harmonicity value of this token interact with the durational constraints.

Cue constraints have to be formalised negatively for the following reason: positively formulated cue constraints fail to predict the correct categorisation in OT if more than two categories are involved, because the highest-ranked cue constraint will always determine the output (for an illustration of this point, see Boersma and Escudero 2008). The advantage of the negatively-formulated cue constraints is that they can also be employed in the modelling of production, where they ensure that a phonological form is realised with the corresponding perceptual cue values, see the bidirectional-phonology-and-phonetics model by Boersma (2006, 2007).

We postulate cue constraints for every possible incoming value along an auditory dimension that is used as a cue in the respective language. But instead of listing large numbers of cue constraints that refer to single values, the present article summarizes these as cue constraints that cover a certain interval along the auditory continuum, and gives them the following form: "x_1 .. $x_2 f$ not /y/" (where x_1 and x_2 are values forming an interval on the auditory continuum f). Furthermore, the simplifying assumption is made that both Dutch and German listeners employ segmental duration and harmonicity median as the only auditory dimensions for labiodentals. In reality, speech perception involves several perceptual cues, which are weighted language-specifically (see e.g. Bradlow 1995). In addition, the present study does not include constraints necessary to identify the segments as labiodentals, though they are of course assumed to apply.[10]

The cue constraints as formalized by Escudero and Boersma (2003) are modified here to map continuous auditory dimensions onto features, and not directly onto segmental categories. By this, we can identify classes that share certain auditory qualities, thereby simplifying the identification of sounds. Furthermore, the mapping of auditory values onto phonological features allows for the normalisation of speech rate (see endnote 9).

The features used in the present study are of a more phonetic nature than those traditionally employed in phonology (e.g., Chomsky and Halle 1968 and Clements 1985). They are not part of a universal set, and thus can differ from language to language, and they do not necessarily provide minimal representations, as we will see below. Instead, the features employed here mirror as closely as possible the phonetic distributions of the classes they describe. This does not prevent their use to describe phonological processes, a point that cannot be illustrated here for lack of space.

The two German labiodentals show a distinction in their duration, which can be mapped onto a binary feature such as [±long].[11] The average production of the Dutch labiodentals in figure 2 shows that all three are (at least partially) differentiated by duration, and therefore cannot be described with a binary feature [±long], alone. Either a ternary feature [long, medium, short] or two binary features [±long] and [±short] have to be used. The present study employs the latter solution for the simple reason that it reduces the number of cue constraints, no theoretical implication hinges on this decision. With respect to the dimension of harmonicity median, the German labiodentals show again a binary distinction, which can be mapped onto the feature [±noise], with /f/ being [+noise] and /v/ [–noise]. The Dutch labiodentals can be argued to show a three-way distinction in harmonicity (though with large overlap of /f/ and /v/), which would require an additional specification with a feature such as [periodic], where /ʋ/ and /v/ are [+periodic], and /f/ is [–periodic]. The resulting feature specifications of the labiodentals are summarised in (4).

(4)

a. *Specification of labiodentals in German*

	/f/	/v/
[long]	+	−
[noise]	+	−

b. *Specification of labiodentals in Dutch*

	/f/	/v/	/ʋ/
[long]	+	−	−
[short]	−	−	+
[noise]	+	+	−
[periodic]	−	+	+

The specifications in (4) are not minimal. The labiodentals in German could be distinguished by either [±long] or [±noise] alone. Similarly, the distinction of the Dutch sounds would require only two features, e.g. [±long] and [±short], or [±long] and [±noise]. Such minimal representations are possible, and actually employed by some speakers, as we will see below. However, they do not cover all perceptual data. The German L1 listeners in Hamann and Sennema's (2005a) perception experiment, for instance, categorised the Dutch stimuli in figure 3 mainly by duration (thus seem to employ a feature [long]), which accounts for a categorisation of /ʋ/ and /v/ as their native /v/ in respectively 99.5 and 82.8 percent of the cases, see table 1 above.

However, further 16.7 percent of Dutch /v/ were categorised as /f/, which is only possible if at least some of the listeners included the dimension of harmonicity (or another measure for noisiness, and thus a feature like [noise]) in their decision. This example illustrates that listeners of the same native language can differ in their use of perceptual cues, and as a consequence, several perception grammars (with different segmental specifications and usage of cue dimensions) are possible for the same language. The specifications proposed in (4) are what is maximally necessary for distinguishing the labiodentals in Dutch and German by duration and harmonicity median.

Based on the feature specifications in (4) and on the average values for the labiodentals in Hamann and Sennema (2005b), the following cue constraints are postulated for German (5a) and Dutch (5b).

(5)

a. *Cue constraints for German*

70 .. 135 ms not [+long]	70 .. 135 ms not [−long]
136 .. 200 ms not [−long]	136 .. 200 ms not [+long]
-6 .. 4 dB not [−noise]	-6 .. 4 dB not [+noise]
5 .. 26 dB not [+noise]	5 .. 26 dB not [−noise]

b. *Cue constraints for Dutch*

70 .. 155 ms not [+long]	70 .. 155 ms not [−long]
156 .. 200 ms not [−long]	156 .. 200 ms not [+long]
70 .. 114 ms not [−short]	70 .. 114 ms not [+short]
115 .. 200 ms not [+short]	115 .. 200 ms not [−short]
-6 .. 8 dB not [−noise]	-6 .. 8 dB not [+noise]
9 .. 26 dB not [+noise]	9 .. 26 dB not [−noise]
-6 .. -1 dB not [−periodic]	-6 .. -1 dB not [+periodic]
0 .. 26 dB not [+periodic]	0 .. 26 dB not [−periodic]

The constraints in the left column of (5) are crucial for the correct classification in both languages and are ranked high, whereas the antagonistic constraints on the right are ranked low. These antagonist cue constraints are not included in the following tableaux for lack of space.

A perception grammar that can correctly categorize the German data from figure 2 above is given in (6). Both tableaux have input auditory forms based on the average values for the German segments obtained in Hamann and

Sennema's (2005b) acoustic study. Candidates are the two phonological segments /f/ and /v/ since these are the only labiodentals in German.

(6) *A German perception grammar*

[v]

[102 ms, 15 dB]	136..200ms not [−long]	70..135ms not [+long]	-6..4dB not [−noise]	5..26dB not [+noise]
/f/		*!		*
☞ /v/				

[f]

[172 ms, -2 dB]	136..200ms not [−long]	70..135ms not [+long]	-6..4dB not [−noise]	5..26dB not [+noise]
/f/				
☞ /v/	*!		*	

As we can see from these tableaux, the ranking among the four constraints is irrelevant for the perception of the German labiodentals, since any ranking will result in the same winning candidates, as long as these four constraints are ranked above their not-included antagonist constraints. Furthermore, we can see that either the two constraints referring to the durational contrast or those referring to the noisiness would be sufficient to decide on the winning candidates. This shows that there is not only one possible perception grammar for German labiodentals but several. These possible grammars differ in the number of cue constraints, depending on which cues the individual listener uses to distinguish the sounds, and/or in the ranking of the constraints, illustrating the different weight individual listeners give to certain cues. We assume the ranking as given in (6), because it predicts the correct naïve perception of Dutch labiodentals (see section 5.1 below).

Similarly, there are several possible perception grammars for the Dutch labiodentals. The most complex one would use all specifications in (4b) and therefore all cue constraints listed in (5b). In (7), we see a perception grammar for Dutch that is less complex and uses only the specifications [±long] and [±noise] and the respective cue constraints, modelling a listener who employs only a binary distinction on the dimensions of duration and harmonicity. The input auditory forms are based on the average values for the Dutch segments from Hamann and Sennema's (2005b) acoustic study. Output candidates are the three labiodental categories in Dutch.

(7) *A Dutch perception grammar*

[ʋ]

[97 ms, 19 dB]	156..200ms not [−long]	70..155ms not [+long]	-6..8dB not [−noise]	9..26dB not [+noise]
/f/		*!		*
/v/				*!
☞ /ʋ/				

[v]

[133 ms, 1 dB]	156..200ms not [−long]	70..155ms not [+long]	-6..8dB not [−noise]	9..26dB not [+noise]
/f/		*!		
☞ /v/				
/ʋ/			*!	

[f]

[170 ms, -2 dB]	156..200ms not [−long]	70..155ms not [+long]	-6..8dB not [−noise]	9..26dB not [+noise]
☞ /f/				
/v/	*!			
/ʋ/	*!			

The use of additional constraints or a different constraint ranking would yield different perception grammars with the same categorisation results. Different results can only occur if the ranking between the constraints in (7) and their antagonists changes, if the feature specification of the segments in (4) is changed, or if we employ different cues than harmonicity and duration.

5. The perception grammars of four types of learners

With the formalisation of the differences in perception grammars between Dutch and German listeners in the previous section, we can now account for the performance of the four types of L2 learners that were introduced in section 2. Since these learners had to categorise the tokens of the speaker given in figure 3, the input to the following tableaux are the average values for these tokens.

Learners of type D do not differ in performance from Dutch speakers and we can conclude that this group has acquired L2 perception grammars identical to that of native listeners. If we assume the Dutch grammar in (7), then the change in input (from the values averaged over 5 female speakers to the average values for the one male speaker) still yields the same results, which is not further illustrated here. Learner types A to C differ in various degrees from native-like performance, as described in detail in the following subsections.

5.1. The L2-perception grammar of Learner A

Learners of type A identified the tokens with /f/ and /ʋ/ correctly and categorised items of the voiced fricative as /ʋ/. This performance is remarkable in as far as the difference in harmonicity median between the tokens of Dutch /v/ and /ʋ/ is not employed as a cue for their categorisation. Learners of type A seem to simply apply their German perception grammar to the Dutch tokens and to equate both Dutch /v/ and /ʋ/ with their native /v/. This indicates that the learners are at an early stage of their acquisition of Dutch and have not created a separate phonological category for a third labiodental in Dutch, yet.

Following Escudero and Boersma (2004) and Escudero (2005), who argue that an L2 learner starts off by copying the native perception grammar, i.e. the phonological categories and cue constraints from the native language (based on the *Full Transfer Theory* by Schwartz and Sprouse 1996), we can assume that learners of type A simply use a copy of their German perception grammar for the Dutch sounds. This is illustrated with the perception tableaux in (8), which contain the constraints and rankings of the German perception grammar in (6). Input are the Dutch tokens, and candidates are the Dutch phonemes /f/ and /ʋ/ (they could also be labelled German /f/ and /v/, respectively). Dutch /v/ is not included, since learners of type A seem to have no segmental representation for this sound.

(8) *Learner A: A German perception grammar (with Dutch labels) on Dutch inputs*

/ʋ/

[120 ms, 22 dB]	136..200ms not [−long]	70..135ms not [+long]	-6..4dB not [−noise]	5..26dB not [+noise]
/f/		*!		*
☞ /ʋ/				

/v/

[125 ms, 3 dB]	136..200ms not [−long]	70..135ms not [+long]	-6..4dB not [−noise]	5..26dB not [+noise]
/f/		*!		
☞ /ʋ/			*	

/f/

[185 ms, -2 dB]	136..200ms not [−long]	70..135ms not [+long]	-6..4dB not [−noise]	5..26dB not [+noise]
☞ /f/				
/ʋ/	*!		*	

The application of a German perception grammar on the Dutch input can account for the performance of naïve monolingual Germans, too. The results of the German L1 listeners in table 1 show that not all tokens of Dutch /v/ were heard as German /v/. 16.7 percent have been categorised as /f/. This is not surprising, if we look at the actual realisations of the tokens in figure 3. Token *va4*, for instance, has a duration of 141 ms which will be categorised by the German perception grammar as /f/:

(9) *A German perception grammar (with Dutch labels) on token "va4"*

va4

[141 ms, 5 dB]	136..200ms not [−long]	70..135ms not [+long]	-6..4dB not [−noise]	5..26dB not [+noise]
☞ /f/				*
/ʋ/	*!			

Token *va1* undergoes a similar misperception as formalised in (9), because it has as a duration of 138 ms, which is categorised as [+long]. All other tokens of /v/ have durations below 136 ms and are therefore categorised as /v/.

The categorisation of some Dutch /v/ tokens as German /f/ could be used to explain why one of the L2 listeners (PP5) categorised only 75 percent of /v/ as Dutch approximant and 25 percent incorrectly. This explanation requires, however, that this listener has separate categories for /f/, /v/ and /ʋ/ and has associated some durational values with /v/. A more likely explanation for this performance is the experimental setup, where three answer categories were provided for the labiodentals, and the participant felt obliged to use the third one (/v/) from time to time, although she does not have three labiodental phonemes.

The second L2 learner that belongs to the group of Learner A showed some additional miscategorisation of /f/ as /v/, a pattern that is covered by type B below.

5.2. The L2-perception grammar of Learner B

Learners of type B perceived the Dutch labiodental approximant /ʋ/ always correctly, but frequently confused the two fricatives. These results can be interpreted in the following way. The learners of type B have only two labiodental categories in their L2 grammar, namely one (the German voiced fricative) for Dutch /ʋ/ and one (the German voiceless fricative) for /v/ and /f/ together. This latter category has two corresponding graphemes, <v> and <f>.[12] Such a grammar is very likely to emerge in an environment where all speakers neutralize the distinction between voiced and voiceless fricatives, something that is reported for a large number of Dutch speakers, recall the description in section 1. As a result, learners receive no input with separate distributions for the labiodental voiced and voiceless fricatives, and thus have no perceptual dimension (neither harmonicity median nor duration) that yields a correct classification, but have to base their decision on chance. We could not control for previous L2 input in our perception experiment, but it seems very likely that some of the learners received only 'neutralising' input.

The learners of type B had to learn to collapse or ignore certain cues and cue values in order to achieve this performance. The cue of duration clearly distinguishes the tokens of /v/ and /f/ in the perception experiment, and therefore the uniform treatment of these tokens can only be attained if duration is not considered as a cue. Furthermore, the tokens of /v/ and /ʋ/ are not distinguished via duration by our speaker, but learners of type B have a clear distinction between these two classes, giving further evidence to the hypothesis that these learners do not employ durational cues. These two observations can be modelled with a perception grammar that does not have constraints on the duration of labiodentals. Furthermore, such a perception grammar does not employ separate representations for /v/ and /f/, see the tableaux in (10).

(10) *A Dutch perception grammar of Learner B*

/ʋ/

[120 ms, 22 dB]	-6..8ms not [−noise]	9..26ms not [+noise]
/f,v/		*!
☞ /ʋ/		

/v/ or /f/

[125..185 ms, -2..3 dB]	-6..8ms not [−noise]	9..26ms not [+noise]
☞ /f,v/		
/ʋ/	*!	

Two of the three learners in group B actually showed some additional miscategorisation of /v/ as /ʋ/. This performance can be explained if duration interacts with harmonicity, and all tokens shorter than for instance 120 ms are categorised as /ʋ/, see the pattern of learner type C described below. The durational difference between German /v/ and /f/ is "unlearnt" for the Dutch perception grammar of these listeners, otherwise we would not find the observed confusion patterns.

An alternative explanation for the learners of type B is that they have created a third category, and distinguish all three labiodentals by harmonicity, only. This requires a specification along the harmonicity dimension with two features such as [±noise] and [±periodicity], recall (4b). Such an account works if we assume as input the data from figure 2, where the categories of /f/ and /v/ are largely overlapping on the harmonicity dimension. The same categories show much less overlap in the tokens in figure 3, however, and we would therefore run into problems when trying to predict the observed confusions with a three-member harmonicity contrast and the input from our male speaker in figure 3.

5.3. The L2-perception grammar of Learner C

Listeners of type C correctly identified Dutch /f/ and categorised some tokens of /v/ as /ʋ/ and vice versa. As the majority of the tokens are classified correctly, we can conclude that these learners have created a third labiodental category for Dutch. A possible explanation for the observed confusions is that tokens of the approximant with a long duration are perceived as /v/, whereas tokens of /v/ with a high harmonicity value are perceived as /ʋ/. Hence both duration and harmonicity are used as cues, and they are traded in a specific relation. Consequently, the resulting perception grammar is more complex than the one formalised in (7) for Dutch listeners. A modelling of this grammar and the cue trading requires a specification of the labiodentals as both [±short] and [±long]. For the features [±short] and [±noise], the boundaries and thus the respective constraints have to be slightly different from those postulated in (4b) and (5b): The learners of type C set the boundary between [+short] and [−short] at 124 ms, not 114 ms, because they categorise tokens with a duration of 114-124 ms as [+short], i.e., /ʋ/. Furthermore, the boundary between [+short] and [−short] lies at 4 dB, not 8 dB, because tokens with harmonicity values between 4 and 8 dB are categorised as [–noise], thus /ʋ/, independent of their length.[13]

To illustrate the perception of such tokens, the input Dutch /ʋ/ is split into two categories in the following grammar; tokens with a duration of 70-124 ms (i.e., /ʋ/ tokens from the five speakers in figure 2), and tokens with a longer duration of 125-155 ms (which are still tokens of /ʋ/ for our male speaker). Both categories have the same harmonicity values of 9-26 dB. Similarly, the input of Dutch /v/ is split into two categories; tokens with a harmonicity value of 1-3 dB and tokens with a harmonicity value of 4-8 dB, both categories having the same durational values.

(11) *A Dutch perception grammar for Learner C*[14]

/ʋ$_1$/

[70..124 ms, 9..26dB]	156..200 ms not [−long]	70..155 ms not [+long]	-6..3dB not [−noise]	70..124 ms not [−short]	125..200 ms not [+short]	4..26dB not [+noise]
/f/		*!		*		*
/v/				*!		*
☞ /ʋ/						

/ʋ$_2$/

[125..155 ms 9..26dB]	156..200 ms not [−long]	70..155 ms not [+long]	-6..3dB not [−noise]	70..124 ms not [−short]	125..200 ms not [+short]	4..26dB not [+noise]
/f/		*!				*
☞ /v/						*
/ʋ/					*!	

/v$_1$/

[70..155 ms, 1..3dB]	156..200 ms not [−long]	70..155 ms not [+long]	-6..3dB not [−noise]	70..124 ms not [−short]	125..200 ms not [+short]	4..26dB not [+noise]
/f/		*!		*		
☞ /v/				(*)		
/ʋ/			*!		(*)	

/v$_2$/

[70..155 ms, 4..8dB]	156..200 ms not [−long]	70..155 ms not [+long]	-6..3dB not [−noise]	70..124 ms not [−short]	125..200 ms not [+short]	4..26dB not [+noise]
/f/		*!		*		*
/v/				*!		*
☞ /ʋ/					(*)	

/f/

[156..200 ms -6..3dB]	156..200 ms not [−long]	70..155 ms not [+long]	-6..3dB not [−noise]	70..124 ms not [−short]	125..200 ms not [+short]	4..26dB not [+noise]
☞ /f/						
/v/	*!					
/ʋ/	*!		*		*	

The actual data from the perception experiment supports the analysis formalised in the perception grammar in (11): A miscategorisation of /v/ as /ʋ/ occurred mainly for tokens of /v/ that had a harmonicity median around 3 dB or above (*va4*: 5.6 dB, *va8*: 3.9 dB, *va3*: 2.9 dB), and there was a tendency to miscategorise /ʋ/ tokens that were rather long (*ʋa4*: 144 ms, *ʋa8*: 122ms, *ʋa1*: 124ms).

Like the native grammar in (7), the learner's grammar in (11) predicts the correct output for the average values of the male speaker in the perception experiment and the average values for the five speakers from the acoustic study. Thus, this grammar is able to deal with prototypical tokens of the three labiodentals, but departs from the L1 grammar in (7) with respect to the categorisation of marginal tokens.

6. Summary and conclusion

The four types of German learners described in section 5 showed very different perceptual strategies for categorising the labiodentals in Dutch. Learner type A simply applied a copy of the German perception grammar, with German representations of the two labiodentals /f/ and /v/ and their respective cue constraints. The learners of type B also had only two labiodental categories for the three Dutch sounds, but had changed their German representations and cue constraints to mirror a Dutch environment that neutralises /f/ and /v/. The learners summarised in Type C were capable of classifying prototypical tokens of the Dutch labiodentals, hence had constructed a category and representations for the voiced fricative and had associated them with the relevant cues. However, these representations and cue constraints proved to be different from those of a native speaker when dealing with non-prototypical tokens. Lastly, learners of type D showed native-like performances, indicating that they acquired representations and cue constraints identical to those of L1 speakers.

One could argue that these four types of learners represent four different stages in the L2 acquisition of Dutch. This is certainly the case for type A and type D, with type A representing the very beginner's stage and type D the end stage. However, there is no evidence for assuming that the increasing correct performance from type B to D indicate three successive developmental stages. The assumption of three successive stages implies that learners first have to have only one representation that combines both labiodental fricatives in Dutch, and employ only harmonicity as a cue to differentiate them from the labiodental approximant. Then they have to move on to construct a third labiodental category and durational cue constraints that distinguish three levels and that interact with harmonicity cues. And lastly they have to get rid again of the third durational category and the interaction of harmonicity and duration constraints.

The present article assumes instead that the perception grammars for types B to D are possible variants of dealing with the same categorisation task, provided the learners had a certain amount of L2 input. These different grammars reflect perception strategies that the learners developed based on their L2 input and on their L1 perception grammar. The role of the received input was elaborated already in section 5.2 for Learner B, who seems to have encountered only neutralised, i.e. voiceless, labiodental fricatives. Learner C and D had received less or no neutralised input, but differ insofar as Learner C probably had a less varied input than Learner D, since Learner C could not deal with all tokens of the Dutch labiodentals provided in the experiment. With respect to the L1 grammar, the present article illustrated that there are several possible perception grammars for German and Dutch that all provide the same categorisation for the native labiodental categories. The idea of learners ending up with different grammars that result in the same output is not new in the OT acquisition literature, see Apoussidou's (2006) work on simulated learning of stress in Pintupi. The role of speaker-specific differences in the L1 grammar for the acquisition of an L2 perception grammar, however, could not be further elaborated here for lack of data. Future research is necessary where detailed perception experiments on L2 and L1 of the same speakers (and possibly also developmental studies on their L2 perception) allow a comparison between the L1 and (several stages of) the L2 grammar.

We saw in the present study that two factors play a role in the account of L2 perception: phonetic factors such as auditory dimensions and the weightings thereof, and phonological factors such as the segmental categories of the native language and their specifications. In this respect, the present study supports earlier L2 studies such as Flege and Hillenbrand (1986) who argue that both phonetic differences between L1 and L2 and the phonology of L1 are crucial in the account of L2 perception. Furthermore, the present study provides a formalisation of how these two factors interact with the help of perception grammars and cue constraints, following Escudero and Boersma's (2003, 2004) groundbreaking work. This enables us to move from a mere description of the data (by comparing phonemes and perceptual cues employed in L1 and L2) to a prediction of L2 perception based on language-specific perception and phonological representations.

The present study is innovative in proposing listener-specific variation in both phonetic and phonological factors of L2 perception: individual choices of auditory dimensions and weighting of these dimensions, as well as individual specifications of phonological contrasts that are not present in the native language.

Notes

1. Parts of the research reported here were presented at the *Phonetik und Phonologie* 2 meeting, Tübingen, July 19, 2005; at the Universities of Amsterdam, Kiel, and Utrecht; and at the ZAS Berlin. I thank the audiences on all of these occasions for their comments. I have also received helpful suggestions and comments by Paul Boersma, Mirjam Ernestus, Frank Kügler, Aditi Lahiri, Anke Sennema, and two anonymous reviewers. I gratefully acknowledge the support of grant GWZ 4/8-1-P2 by the German Science Foundation (DFG) and a VENI postdoctoral fellowship 16.064.057 by the Dutch Science Foundation (NWO).
2. Whereas the symbol /β/ stands for a bilabial fricative, the underscore /˕/ indicates that it is lowered, i.e., an approximant. The symbol /w/ is sometimes also used for the Southern Dutch/Flemish sound (e.g. Cohen et al. 1961), but incorrectly implies a secondary velarisation.
3. German has the grapheme <v>, which is used both for /v/ and /f/, see e.g. *Vase* [vaːzə] 'vase' and *Vieh* [fiː] 'cattle', respectively.
4. The Dutch alveolo-palatal sibilant [ɕ] is usually considered an allophone of /s/, see Mees and Collins (1982: 2).
5. A statistical analysis of the data that would yield a clustering of speakers is unfortunately not possible. The most appropriate statistical method for this purpose is a generalised linear mixed-effects model (Lindstrom and Bates 1990). However, at the time this paper was finished such a model could not yet deal well with categorical dependent variables.
6. 5. The performance of these two participants is simplified here for reasons that become obvious in §5.1: PP11 showed some additional miscategorisation of /f/ as /v/, a pattern that is covered by learner type B. PP5 categorized 25% of /v/ incorrectly, as accounted for in §5.1.

7. PP20 and PP21 actually categorize /v/ as /ʋ/ in 30 percent of the cases. This confusion pattern is covered by Learner C.
8. The speaker that was used in the perception experiment did not distinguish /v/ and /ʋ/ by duration (cf, figure 3), and the German listeners therefore could not employ duration alone for categorizing these sounds in the experiment. Since the average distributions of the three labiodentals in figure 2 (from the five speakers in Hamann and Sennema 2005b) show that there is large overlap of the tokens of all *three* categories along this dimension, duration alone does not seem to be a reliable cue for the Dutch labiodentals in general.
9. Normalisation for speaking rate is possible if we employ durational cue constraints that map auditory values onto phonological features such as [±long], see (5). This allows a relativization of values according to overall speech rate. The actual values of a constraint like "x–y ms is not [+long]" in the perception of a labiodental, for instance, can be determined by the duration of adjacent vowels or the larger context (word, utterance) in which the segment appears. Similar normalization of speaker-dependent differences such as differences in formant values can be performed via contextual information (like the pitch range within the word or utterance).
10. Cues employed for identifying place of articulation are for instance the spectral features of the friction noise (e.g. Whalen 1991) and vowel transitions (e.g. .Bladon et al. 1987; see also Nowak 2006 who illustrates the interaction of both cues in the perception of the Polish sibilants).
11. A larger contrast on the durational dimension might be necessary for other than labiodental contrasts in German, see e.g. Hamann's (2003) proposal to account for the difference between long tense vowels, short tense vowels and glides of the same place of articulation in German, such as /iː, i, j/, by a three-way durational contrast.
12. This situation is comparable to German, where <f> corresponds to /f/, and both <v> and <w> correspond to /v/, see endnote 3.
13. As the constraints employed in this article cover certain intervals on auditory dimensions, the change of boundary values actually means the following. The high-ranked "115 ms not [+short]", which was integrated in the constraint "115..200 ms not [+short]", is now ranked lower than its antagonist "115 ms not [–short]", which is now integrated in "70..124 ms not [–short]". The same holds for the durational constraints for 116-124 ms. The change is thus one in ranking, not in values. This also applies to the 'boundary' change of the harmonicity constraints.
14. The violation marks in brackets indicate that the respective constraints are violated by part of the input values. Due to the low ranking of these constraints, the input is not further split to illustrate these violations.

References

Apoussidou, Diana
2006 Learning to be insensitive to weight in Pintupi. Ms., ROA 799.

Best, Catherine T., and Winifred Strange
1992 Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20: 305-330.

Bladon, Anthony, Christopher Clark, and Katrina Mickey
1987 Production and perception of sibilant fricatives: Shona data. *Journal of the International Phonetic Association* 17(1): 39-65.

Boersma, Paul
1993 Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proceedings of the Institute of Phonetic Sciences (University of Amsterdam)* 17: 97-110.

Boersma, Paul
1998 *Functional Phonology: Formalizing the Interaction Between Articulatory and Perceptual Drives.* The Hague: Holland Academic Graphics.

Boersma, Paul
2006 Prototypicality judgments as inverted perception. In *Gradience in Grammar*, Gisbert Fanselow, Caroline Féry, Ralf Vogel and Matthias Schlesewsky (eds.), 167-184. Oxford: Oxford University Press.

Boersma, Paul
2007 Some listener-oriented accounts of h-aspiré in French. *Lingua* 117: 1989-2054.

Boersma, Paul, and Paola Escudero
2008 Learning to perceive a smaller L2 vowel inventory: An Optimality Theory account. In *Contrast in phonology: theory, perception, acquisition*, Peter Avery, Elain Dresher and Keren Rice (eds.), 271-301. Berlin and New York: Mouton de Gruyter.

Booij, Geert
1995 *The Phonology of Dutch.* Oxford: Oxford University Press.

Bradlow, Ann R.
1995 A comparative study of English and Spanish vowels. *Journal of the Acoustical Society of America* 97: 1916-1924.

Chomsky, Noam, and Morris Halle
1968 *The Sound Pattern of English.* New York: Harper and Row.

Clements, George N.
1985 The geometry of phonological features. *Phonology Yearbook* 2: 225-252.

Cohen, A., C.L. Ebeling, K. Fokkema, and A.G.F. van Holk
1961 *Fonologie van het Nederlands en het Fries.* 'S-Gravenhage: Martinus Nijhoff.

Escudero, Paola
2005 *Linguistic perception and Second Language Acquistion: Explaining the attainment of optimal phonological categorization.* Utrecht: LOT.

Escudero, Paola, and Paul Boersma
2003 Modelling the perceptual development of phonological contrasts with Optimality Theory and the Gradual Learning Algorithm. In *Proceedings of the 25th Annual Penn Linguistics Colloquium (Penn Working Papers in Linguistics 8.1)*, Sundha Arunachalam, Elsi Kaiser and Alexander Williams (eds.), 71-85.

Escudero, Paola, and Paul Boersma
2004 Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26: 551-585.

Faulkner, Andrew, and Stuart Rosen
1999 Contributions of temporal encodings of voicing, voicelessness, fundamental frequency, and amplitude variation to audio-visual and auditory speech perception. *Journal of the Acoustical Society of America* 106(4): 2063-2073.

Flege, James
2002 Interactions betwen the native and second-language phonetic systems. In *An integrated view of language development: Papers in honnor of Henning Wode*, P. Burmeister, Thorsten Piske and A. Rohde (eds.). Trier: Wissenschaftlicher Verlag.

Flege, James Emil, and James Hillenbrand
1986 Differential use of temporal cues to the /s/–/z/ contrast by native and non-native speakers of English. *Journal of the Acoustical Society of America* 79(2): 508-517.

Gussenhoven, Carlos
1999 Illustrations of the IPA: Dutch. In *Handbook of the International Phonetic Association* 74-77. Cambridge: Cambridge University Press.

Hamann, Silke
2003 German glide formation functionally viewed. *ZAS Working Papers in Linguistics* 32: 137-154.

Hamann, Silke, and Anke Sennema
2005a Voiced labiodental fricatives or glides - all the same to Germans? In *Proceedings of the Conference on Plasticity in Speech Processing*, Valery Hazan and Paul Iverson (eds.), 164-167. London: UCL.

Hamann, Silke, and Anke Sennema
2005b Acoustic differences between German and Dutch labiodentals. *ZAS Working Papers in Linguistics* 42: 33-41.

Hayes, Rachel
2002 The perception of novel phoneme contrasts in a second language: A development study of native speakers of English learning Japanese

singleton and geminate consonant contrasts. In *University of Arizona Coyote Working Papers 12: Language in Cognitive Science*, Rachel Hayes, W.D. Lewis, E.L. O'Bryan and Tania Zamuner (eds.), 28-41. Tuscon: University of Arizona Press.

Iverson, Paul, Patricia K. Kuhl, Reiko Akahane-Yamada, Eugen Diesch, Yoh'ich Tohkura, Andreas Kettermann, and Claudia Siebert
2003 A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* 87: 47-57.

Jessen, Michael
1998 *Phonetics and phonology of tense and lax obstruents in German.* Amsterdam: Benjamins.

Kohler, Klaus
1995 *Einführung in die Phonetik des Deutschen*. Berlin: Erich Schmidt.

Kohler, Klaus
1999 Illustrations of the IPA: German. In *Handbook of the International Phonetic Association* 86-89. Cambridge: Cambridge University Press.

Lindstrom, Mary J., and Douglas M. Bates
1990 Nonlinear Mixed Effects Models for Repeated Measures Data. *Biometrics* 46(3): 673-687.

McCarthy, John J., and Alan Prince
1993 Prosodic morphology I: Constraint interaction and satisfaction. Ms., *University of Massachusetts and Rutgers University.*

Mees, Inger, and Beverly Collins
1982 A phonetic description of the consonant system of Standard Dutch (ABN). *Journal of the International Phonetic Association* 12(1): 2-12.

Nowak, Pawel M.
2006 The role of vowel transitions and frication noise in the perception of Polish sibilants. *Journal of Phonetics* 43(2): 139-152.

Pater, Joe
2004 Bridging the gap between perception and production with minimally violable constraints. In *Constraints in phonological acquisition*, René Kager, Joe Pater and Wim Zonneveld (eds.). Cambridge: Cambridge University Press.

Plack, Christopher J.
2004 Auditory Perception. In *Psychology: An international perspective*, Michael W. Eysenck (ed.) 1-19. New York: Psychology Press.

Prince, Alan, and Paul Smolensky
1993 Optimality Theory: Constraint Interaction in Generative Grammar. Ms., *Rutgers University*. Brunswick.

Schwartz, Bonny, and R.A. Sprouse
1996 L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 12: 40-72.

Stevens, Kenneth N., Sheila Blumstein, Laura B. Glicksman, M. Burton, and Kathleen Kurowski
1992 Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of the Acoustical Society of America* 91: 2979-3000.

Viëtor, Wilhelm
1897 *Elemente der Phonetik des Deutschen, Englischen und Französischen.* Leipzig: Reisland (5th edition 1904).

Wängler, Hans-Heinrich
1974 *Grundriss einer Phonetik des Deutschen.* Marburg: N.G. Elwert Verlag.

Whalen, D. H.
1991 Perception of the English /s/-/ʃ/ distinction relies on fricative noises and transitions, not on brief spectral slices. *Journal of the Acoustical Society of America* 90(4): 1776-1785.

Wiese, Richard
1996 *Phonology of German.* Oxford: Oxford University Press.

Prosodic conditioning, vowel dynamics and sound change

Ewa Jacewicz, Joseph Salmons & Robert Allen Fox

1. Introduction

This paper reports on research aimed at bridging two important but disparate approaches to understanding variability and change in the realization of vowels: (1) phonetic work on how prosody shapes vowel characteristics and (2) diachronic and sociolinguistic work on ‘ chain shifting’ in vowel systems. We will examine two types of phonetic variation: changes in the dynamic formant patterns of vowels as a function of levels of prosodic emphasis and phonetic quality differences associated with regional dialects of American English. Understanding the nature and origins of such phonetic variation will, we believe, provide insight into the sources of sound change over time.

With production and perception data from two Midwestern dialects, we explore parallels between the effects of prosodic emphasis on vowel characteristics and the general directions of change over time and across languages observed for chain shifts. Those parallels shed light on a conundrum known as the ‘perseverance problem’ in chain shifts (Stockwell 1978), namely how generation after generation of speakers could continue to move vowels slowly but consistently in particular directions. The teleological notion that generations are all working toward the same goal in such changes must be rejected, and hypothesized social motivations, while promising, are incomplete. Furthermore, none of this helps to explain why similar shifts occur time and again across languages and periods. One possible structural explanation, for which a foundation is proposed in this paper, is that emphatic realizations cause some vowels to be perceived as more prominent (and thus more salient) in the continuously varying speech stream. These more emphatic productions may serve as the primary impetus for vowel shift over time in that more forceful realizations of vowels are transmitted across generations as less marked ones.

We see variation in vowel enhancement as interacting with vowel properties to bring about a shift as language is handed on from one generation to the next. As vowels are acquired by new generations, the older generation’s more emphatic renditions map to less prominent realizations in the speech of younger speakers, yielding gradual shift-like patterns of change. We illustrate this informally below, where the interlocking circles represent the varying realizations of the tense vowel /e/ under the relevant degrees of stress, with an older generation's prominent realizations mapping to a younger generation's prototypical form.

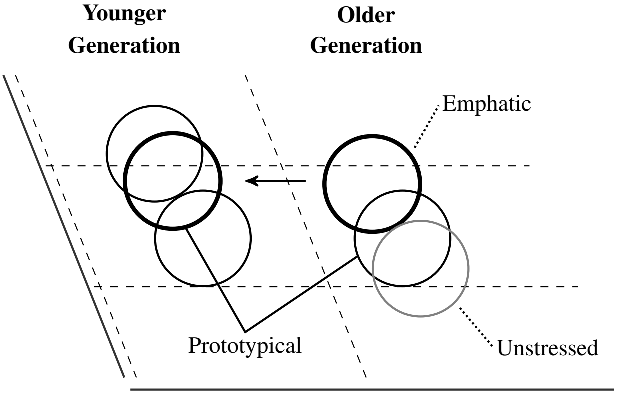


Figure 1: Schematic representation of chain shifts as a function of graded emphasis (or prominence) in vowel production involving two generations of speakers.

1.1. Regional differences in American English vowels and principles of vowel chain shifts

A body of work since Labov, Yeager and Steiner (1972) through Labov (1994, 2001) and Labov et al. (2006) argues that urban areas from Buffalo, New York to Madison, Wisconsin are undergoing a new chain shift (the Northern Cities Shift), including the raising of the vowel in words like *bad* and *bag*, which as a result can sound to speakers from other areas like *bed* and *beg* or even *bid* and *big*. In the Southeast, in contrast, lax front vowels, as in *bit*, are tensing, lengthening and even diphthongizing, so that *bit* can be pronounced like [biːjɪt], while tense front vowels, like in *beat*, tend to become laxer and more centralized. In the West, we find mergers notably of the vowels in *cot* and *caught*, so that *Don* and *Dawn* now sound identical for many speakers. Across much of the U.S. – outside of the Upper Midwest – the tense high back vowel /uː/ is fronting, and its mid counterpart /oː/ is fronting in the south and east. These and similar changes constitute a profound and rapid regional diversification of everyday American English speech.

These differences define the core of the emerging field of sociophonetics, and are central topics in sociolinguistics, dialectology and historical linguistics. Most theoretical discussions in the literature are organized around the notion of *chain shifts*, the tendency, in at least some languages, for individual vowels not to change in isolation from each other over time, but rather to move apparently linked together through the traditional vowel space. Such patterns have long been recognized, most famously in the so-called *Great Vowel Shift* of English, and the vast literature on chain shifts has been reviewed recently by Labov (1994, 2001), Gordon (2001, 2002) and others. A consistent pattern of chain shifting has been posited since at least the German phonetician Sievers (1876/1881) and the core observation has been often repeated down to the present, as by Labov (1994: 116, drawing on his "preliminary formulation"):

(1) Principle I: In chain shifts, long (tense) vowels rise.
Principle II: In chain shifts, short (lax) vowels fall.

These patterns are widespread and chronic in the Germanic languages, along with some other families, to the point that it has been argued that instead of talking about particular chain shifts as happening at particular times, they should be regarded as an omnipresent, dynamic characteristic of vowels in a language like English (Stockwell 1978). When compared closely to the historical record of attested chain shifts, Principle I is broadly and securely attested – Labov (1994: 122, and elsewhere) alone gives 15 examples of long vowels rising in apparent chains. But Principle II is relatively ill-attested (Labov gives only two examples), and short vowels clearly show far more complex and variable patterns of change than this principle predicts. Take, for example, the already-mentioned Northern Cities Shift, currently in progress across urban areas from upstate New York westward to Wisconsin. In this shift /ɪ/ is generally lowering to /ɛ/, but /ɛ/ is generally thought to be backing to /ʌ/. At the same time, some studies like Gordon (2001) show widespread variation here. For example, /ɛ/ lowers or backs, while /ʌ/ backs, raises or lowers. We conclude that chain shifting of /ɛ/ typically involves lowering and/or backing.[1]

The present study investigates changes to the acoustic characteristics of vowels (including durational and spectral changes) spoken with graded degrees of emphasis in two closely related regional dialects of Midwestern American English, central Ohio and southern Wisconsin. Wisconsin English is assumed by sociolinguists to be in an early stage of the Northern Cities Shift whereas English spoken in central Ohio is not currently participating in any chain shift. It is therefore of interest whether changes to the acoustic structure of vowels caused by the variation in vowel enhancement are different in these two dialects. If diachronic principles are rooted in synchronic variation, we would expect long vowels to rise in the acoustic space in most prominent or emphatic productions. Although the extent of this raising in each dialect, as well as perceptual responses to such raising effects, are difficult to predict on the basis of broadly defined diachronic principles, we would generally expect more enhanced short vowels to lower and/or back in the acoustic space although changes in progress related to Northern Cities Shift in Wisconsin English may somewhat obscure these effects. The study considers acoustic and perception data in the search for synchronic evidence which could serve as a basis for explanation of diachronic principles of sound change.

2. The acoustic study

We examine the nature of acoustic variation that results from a specific prominence-defined position of a vowel in an utterance. We predict that vowels in the most prominent or emphatic positions in an utterance will be acoustically more enhanced than vowels in the positions of intermediate prominence, but the latter will still show more of such enhancement than vowels in the weakest position. In this paper, we refer to the position of highest prominence as the Utterance-initial position (U) followed by the Phrase position (P) and Syllable position (S), respectively.[2] The acoustic measures employed here include vowel duration and the frequencies of the first two formants, F1 and F2. Formant frequencies are measured at five different temporal points in the vowel to allow an approximation of formant movement in the course of vowel's duration. The prediction is that vowels in U-position are longer and their formant values are more extreme as compared to P- and S-positions, respectively.

2.1. Method

2.1.1. *Speakers*

Sixteen native speakers of two Midwestern varieties of American English participated in the experiment. Eight speakers (four males and four females) were born and raised in the Columbus, Ohio area, and eight speakers (four males and four females) were born and raised in the area of Madison, Wisconsin. Their ages ranged from 16 to 31 at the time of recording. All speakers were students, either high school students (Columbus) or undergraduate and graduate students of various majors enrolled at The Ohio State University or University of Wisconsin-Madison. The speakers were unaware of the purpose of the experiment.

2.1.2. *Stimulus materials*

The long vowel /e/ and the short vowel /ɛ/ were selected and placed in existing monosyllabic words in /b_t/ environments. Speakers produced the following sentences containing the words *bait/bet* along with distractor sentences in three distinct positions of prosodic prominence, here labeled U, P, S:

(2) *The strongest U-position (representing the stressed syllable in a noun phrase in utterance-initial position)*
‘Bait shop’ is what I said.
‘Bet some’ is what I said.

(3) *The intermediate P-position (representing the stressed syllable in a noun-phrase in non-initial position in an utterance)*
She said the bait shop was closed.
He said the bet slips were here.

(4) *The weakest S-position (representing the unstressed but unreduced syllable in a morphological compound)*
Shark bait with flavor seems hard to find.
Risky bets are nice but safe bets are better.

These three different sentence structures were used to elicit the productions of [beɪt] and [bɛt] with consistently different degrees of prosodic prominence, phonetically reflected in graded differences in emphasis. We make no theoretical claims regarding the phonological/prosodic structure of these phrases.[3]

2.1.3. *Procedure*

Each speaker read one sentence at a time in a random order for a total of three repetitions. Recording, procedure, and initial data processing were controlled by a program written in MATLAB. The sentences were recorded directly onto a hard drive at a 44.1-kHz sample rate while the speaker was seated in a sound-treated IAC booth. A head-mounted microphone (Shure SM10A) was used, placed at a 1-inch distance from the lips. Each speaker was instructed to read a sentence appearing on the monitor screen in a way typical of his/her conversational speaking style ("as you normally say it while talking to someone"). No other specific reading instructions were given to the speakers.

2.1.4. *Vowel measurements*

Vowel duration and formant movement across vowel duration served as primary measures of the effect of prosodic prominence on the acoustic changes in vowels. Vowel duration was measured from waveform with reference to a spectrogram, using TF32 speech analysis software (Milenkovic 2003). Vowel onsets and offsets included formant transitions. The initial measurement point was located at the first positive peak in the periodic waveform and the final measurement location was at the beginning of the stop closure. All segmentation decisions were later checked and corrected (and then re-checked) using a MATLAB program that displayed the segmentation marks superimposed over a display of the token's waveform. The frequencies of the first two formants of each vowel were then obtained at positions corresponding to 20%, 35%, 50%, 65% and 80% of duration of the vowel. These five formant measurements should exclude immediate consonant influences during CV or VC formant transitions. The stimulus tokens were downsampled to 11.025 kHz and premphasized (98%) prior to spectral analysis. Formant frequencies were estimated with a MATLAB program that utilized a 14-order LPC analysis with a 15 msec Hanning window centered over each measurement location.

2.2. Results

2.2.1. *Vowel duration*

Mean vowel durations are shown in Figure 2. As is evident, for both the long vowel /e/ and the short vowel /ɛ/ duration was sensitive to the degree of vowel enhancement. Higher prominence was reflected in longer durations, particularly in the strongest U-position. The differences between intermediate and weak positions were small but also consistent with the graded prosodic prominence except of /e/ in the S-position in Wisconsin productions. The dialectal differences were manifested in longer durations of vowels spoken by Ohio speakers.

To assess the significance of these results, a mixed design analysis of variance (ANOVA) was conducted (using the SPSS v. 13 software package) with the within-subject factor prosodic position (U, P, S) and the between-subject factors speaker gender (male, female) and speaker dialect (Ohio, Wisconsin). For all reported within-subjects significant main effects and interactions, the degrees of freedom were Greenhouse-Geisser adjusted when there were significant violations of sphericity. Post hoc analyses, when reported—both for vowel duration and formant analyses—were completed using additional ANOVAs on selected subsets of the data (with appropriate F-tests) and either the Tukey test (for between-subject factors) or GLM contrasts (for within-subject factors).

For the vowel /e/, there was a significant main effect of prosodic position ($F(1.7, 20.8) = 13.64$, $p < 0.001$, $\eta^2 = 0.532$) and post-hoc tests showed that, for each dialect, vowels in the U-position were significantly longer than vowels in either the P- and S-position, which did not differ significantly from each other. Thus, the observed graded differences between the P- and S-positions in the Ohio productions were not large enough to reach statistical significance. The effect of speaker dialect was not significant although the Ohio speakers produced longer vowels than the Wisconsin speakers due to the dialectal differences in the U- and P-positions. Similarly, although the effect of gender was not significant, the mean vowel durations were longer for the female speakers than for the male both in Ohio and Wisconsin.

For the vowel /ɛ/, the main effect of prosodic position was significant ($F(1.5, 18.2) = 9.4$, $p = 0.003$, $\eta^2 = 0.440$), with significantly longer vowels produced in the U- position than in either the P- or S-positions. This result is consistent with that for the vowel /e/, indicating that the graded differences

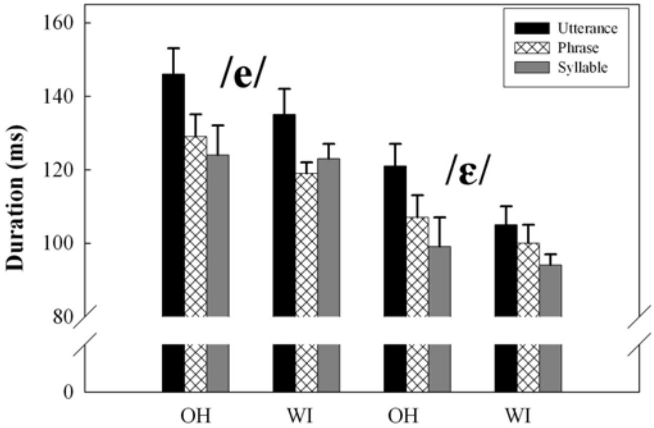


Figure 2: The effect of prosodic prominence on the duration of the vowels /e/ and /ɛ/ in the production of Ohio and Wisconsin speakers. The bar graphs represent mean values. Error bars indicate one standard error.

between the P- or S-positions for either dialect were not large enough to reach significance. Vowels spoken by Ohio speakers were again longer than vowels in Wisconsin productions although these differences were not significant. Speaker gender was significant for /ɛ/ ($F(1, 12) = 6.9$, $p = 0.022$, $\eta^2 = 0.364$), indicating that female vowels were longer across all prosodic positions than male vowels.

Overall, the duration results show that vowels in the highest prominence position were significantly longer than those produced in less prominent positions. At present, the remaining differences found in the means displayed in Figure 2, although suggestive with regard to graded effects of prosodic hierarchy, should be regarded as tendencies rather than representing a conclusive statement.

2.2.2. *Vowel formant pattern*

Figure 3 shows, for each prosodic prominence position, the average changes in F1 and F2 frequencies at five locations in the vowel. As can be seen, the specific formant frequency patterns for either vowel /e/ or /ɛ/ are different in the three prosodic positions. Comparing the locations of the vowels in the acoustic vowel space across the three prosodic positions, one can observe that the trend of greater *raising* and fronting of the vowel /e/ is associated with increased prosodic prominence. A similar effect is also evident for the vowel /ɛ/ although more extreme formant frequency values caused by more emphatic vowel productions corresponded to *lowering* and fronting of the vowel with increased prosodic prominence. As can be seen, frequency changes for F2 for /ɛ/ are small as compared to /e/ due to the non-diphthongal nature of this vowel.

As determined by a series of repeated-measures ANOVAs, enhancement due to the prosodic context had a significant effect on the frequencies of both F1 and F2 in the production of /e/ for both Ohio and Wisconsin speakers (see Table 1 in the Appendix for details concerning ANOVA results). Exploring the significant main effect of prosodic prominence on F1 changes, post-hoc analyses showed that the overall frequencies of F1 were significantly lower in the U-position than in the P-position, and the latter were still significantly lower than in the S-position. Consistently, the overall frequencies of F2 were significantly higher in the U-position than in the P-position, and higher in the P-position than in the S-position. These results are consistent with the graded positions of prominence and show that the degree in vowel enhancement influences changes in formant frequencies over the duration of the vowel in a predictable way.

In addition to the main effect of prosodic context, a significant interaction between prosodic position and vowel measurement location (except for F1 in the Wisconsin productions) showed that prosodic position systematically influenced the specific pattern of frequency change over the course of the vowel's duration. In the most prominent U-position, the vowels exhibited more diphthongal qualities in terms of the extent of frequency change (both F1 and F2) over their durations than vowels in P- and S-positions, respectively.

For the vowel /ɛ/, prosodic context had an effect on both formant frequency values and the nature of the frequency change in the course of the vowel's duration (see Table 2 in the Appendix for details concerning ANOVA results). The significant main effect of prosodic position for F1 for both Ohio and Wisconsin speakers revealed that the vowels in S-position had significantly lower overall mean frequencies, which corresponds to their raised positions in the acoustic space relative to the two more prominent prosodic positions. The effect of prosodic position was also significant for F2. The F2 values for /ɛ/ indicated that the vowel was gradually fronted in the course of its duration for the more prominent U- and P-positions and it was centralized

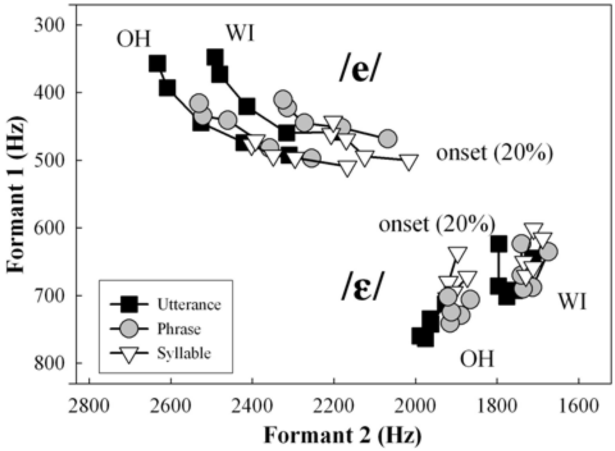


Figure 3: Mean F1 and F2 frequencies for /e/ and /ɛ/ at five temporal locations in the vowel, across prosodic positions. The 20% measurement location is marked as 'onset' and the lines connect each consecutive measurement for each prosodic level.

in the least prominent S-position. This directional change in F2 resulting from an interaction of prosodic context and a particular vowel measurement location could not be detected if F2 values were measured at one location only.

With regard to dialectal differences displayed in Figure 3, Wisconsin vowel productions appear to be more centralized in comparison with Ohio vowels. Because we do not have measures of entire vowel spaces for Ohio and Wisconsin at present, we do not know whether the dialectal differences include the shift of an entire vowel space toward the back for Wisconsin speakers. Based on the present data only, we see that both vowels /e/ and /ɛ/ are more fronted in Ohio relative to vowels spoken in Wisconsin.

2.2.3. *Vector length*

Analyses of F1 and F2 measures, separately, provide indirect insight into the total trajectory of vowel movement in the F1 by F2 plane. In order to examine the magnitude of this movement, we utilized the measure "vector length" or VL (Ferguson and Kewley-Port 2002, also see Hillenbrand et al. 1995) which is the unsigned linear distance (in Hz) between the start of the vowel (i.e., the location of the 20% point) and the end of the vowel (the 80% point) in the F1 by F2 plane. Diphthongal or diphthongized vowels will have longer vector lengths than will monophthongs, which represent their greater amount of frequency change. Consequently, we expect that vowels in more prominent prosodic positions will have longer vector lengths because of the greater F1 and F2 frequency change associated with increased prosodic prominence.

Figure 4 shows the average VL differences for the vowel /e/ as a function of prosodic context. For the Ohio speakers, the VL of the vowels varied significantly by prosodic position ($F(1.9,\ 11.3) = 40.93$, $p < 0.001$, $\eta^2 = 0.872$), with the U-, P- and S-positions each differing significantly from the other. Longer VLs were found in the more prominent prosodic positions. The Wisconsin speakers exhibited similar patterns for VL. There was a significant prosodic context effect ($F(1.2,\ 7.3) = 12.76$, $p = 0.007$, $\eta^2 = 0.680$) with the U-, P- and S-positions differing significantly one from the other. As for the Ohio speakers, longer VLs were associated with greater prosodic prominence. The longer vectors reflected more diphthongal vowel characteristics, which significantly increased from the least prominent prosodic position to the most prominent, consistent with the prosodic hierarchy.

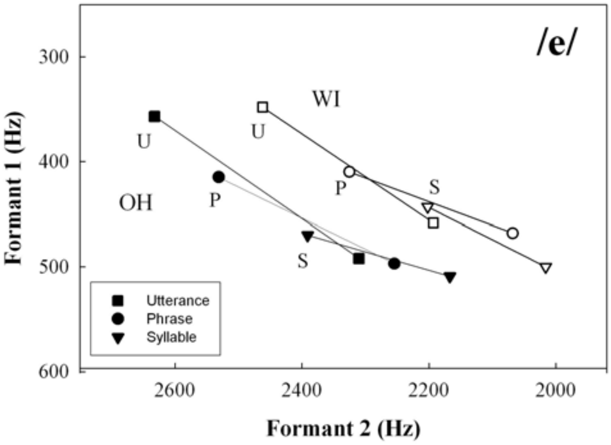


Figure 4: Change of F1 and F2 frequency between the 20% and 80% measurement points for the vowel /e/, representing the 'vector length' or the amount of frequency change across prosodic positions. The symbols identifying the prosodic level (U, P, S) are plotted at the 80% location and a line connects this point to the 20% location.

Based on a visual comparison of VLs for Ohio and Wisconsin vowels in Figure 4, one can observe that VLs in both U- and P-positions are longer for Ohio vowels. This indicates that the vowel /e/ spoken in Ohio is more diphthongal compared to that spoken in Wisconsin, at least in more prominent prosodic positions.

As might be expected, VLs for the monophthongal /ɛ/ were, in general, smaller than those for /e/ for both Ohio and Wisconsin speakers but there were some indications that prosodic context affected VLs for /ɛ/ as well (see Figure 5). However, these effects were not as strong and somewhat mixed in terms of directionality of the frequency change across prosodic positions. For both Ohio and Wisconsin speakers, ANOVA of the VL data yielded no significant effect as a function of prosodic context. As shown in Figure 5, there is much more variation in terms of VL differences between the two dialects, particularly for the P- and S-positions.

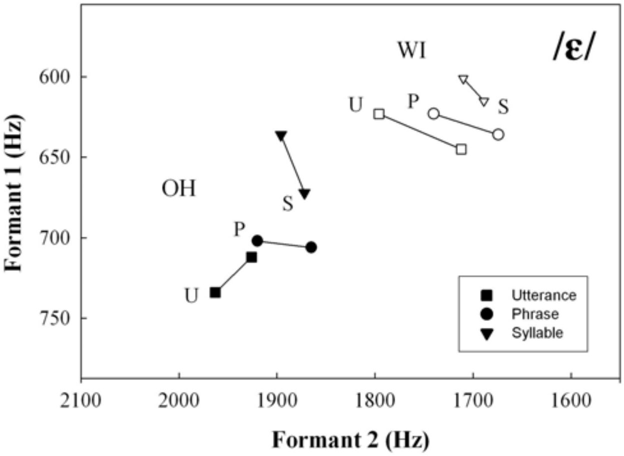


Figure 5: Change of F1 and F2 frequency between the 20% and 80% measurement points for the vowel /ɛ/, representing the 'vector length' or the amount of frequency change across prosodic positions. The symbols identifying the prosodic level (U, P, S) are plotted at the 80% location and a line connects this point to the 20% location.

3. Discussion of the acoustic study

Investigating differences in vowel enhancement as one source of systematic variation in vowels, this study shows that greater prosodic prominence can produce notable changes to both vowel duration and the frequencies of formants over the duration of a vowel. These changes result in progressive raising of /e/ and progressive lowering of /ɛ/ in the acoustic vowel space with each higher prominence position.

Considering duration, vowels in the prosodically strongest U-position had significantly longer durations than in either P- or S-positions. This was true both for the tense vowel /e/ and the lax vowel /ɛ/. The lack of significant differences between the two less prominent prosodic positions may indicate that duration alone is not a strong indicator of prosodic hierarchy despite seemingly straightforward predictions about temporal differences to be found across prosodic levels. It is worth noting that results of other studies are even less conclusive in detecting the relation between prosodic strengthening and vowel duration than the present data, including languages with considerably different accentual/prosodic systems. For example, Onaka (2003) reported no stable pattern of vowel duration in postboundary position (V2) across four prosodic domains for the two Japanese subjects in the study. In another study, Cho and Keating (2001) reported mixed results in the duration of V2 across prosodic domains in Korean. The differences in vowel duration as a function of prosodic position were statistically not significant and the direction of variation differed depending on the domain-initial consonant.

Although not significant, the dialectal differences in vowel duration between the Ohio and Wisconsin variants showed a systematic pattern, which has also been found in another study conducted with different speakers from these two dialects (Jacewicz et al. 2007). Namely, under comparable experimental conditions, Ohio speakers produced longer vowels than Wisconsin speakers. These results were rather unexpected as such cross-dialect differences in vowel duration have not been previously reported in American dialects, although this topic warrants further investigation.

Examining the formant pattern, we measured the frequencies of F1 and F2 at five different locations over the duration of each vowel to capture the dynamic nature and extent of frequency change and its potential variations resulting from prosodic context. This technique proved successful in tracking variation in formant frequencies at different prosodic positions. Measured in this way, the frequency data revealed strong significant effects of prosodic position, measurement location and, most importantly, the interaction between these two which indicated a progressive change in location of the vowel in the acoustic space under greater prominence such as raising and fronting of /e/ and lowering and fronting of /ɛ/.

The vector length analysis further assessed the amount of frequency change over each vowel's duration, showing that vowels in higher prosodic prominence positions become more diphthongal than vowels in lower positions. These changes were also evident for the monophthongal /ɛ/, especially in the Wisconsin data, although generally the vectors were much shorter as compared to the diphthongal vowel /e/.

Relating our acoustic results to diachronic vowel changes, we see strong parallels to the behavior reported for long/tense vowels in chain shifts. Most importantly, the vowel /e/ rises under greater prosodic prominence, in step with diachronic Principle I. For lax (or short) /ɛ/, the literature on chain shifts allows less firm predictions. Still, there are clear parallels, notably that we find lowering of /ɛ/ with greater prosodic prominence, in line with diachronic Principle II. The present results demonstrate ways in which both of these shifts might occur, suggesting that prosodically structured variation in vowels is one motivation for the direction of movement in chain shifts.

Although we expected similar acoustic effects as a function of vowel enhancement for both Ohio and Wisconsin productions, we also expected some differences related to the fact that Wisconsin vowels participate in the Northern Cities Shift whereas vowels spoken in central Ohio do not undergo any shift. We found Ohio vowels to be longer and more fronted relative to Wisconsin productions. In addition, the tense /e/ appears to be more diphthongal. These findings allowed us to formulate specific predictions about the perception of these dialectal differences in vowel characteristics.

For /e/, we can predict that the most emphatic productions in the positions of highest prominence will yield the highest identification of the vowel for Ohio productions but not for vowels spoken by Wisconsin speakers. This is because Ohio vowels represent clearer exemplars of the category relative to Wisconsin vowels, being longer, more fronted, and more diphthongal. Conversely, Wisconsin vowels may sound more monophthongal relative to Ohio vowels, which may result in confusions with neighboring monophthongs /i/, /ɪ/, or /ɛ/.

For the short /ɛ/, we predict more confusions with neighboring short vowels, particularly in the lowest prominence position. Since Ohio /ɛ/ is more fronted than Wisconsin /ɛ/, we would expect the vowel to sound more like an

/æ/ in Ohio productions but not as much when spoken by Wisconsin speakers. Because of its more centralized position in the acoustic space relative to Ohio vowels, Wisconsin /ɛ/ may be perceived more often as /ɪ/ or /ʌ/.

4. The perception study

A vowel identification experiment was conducted to verify our predictions about perceptual responses. Listeners from either Ohio or Wisconsin listened to the productions of both Ohio and Wisconsin speakers. For each dialect, listeners and speakers came from the same geographic areas.

4.1. Stimuli, listeners, and procedure

Stimuli consisted of all instances of the words "bait" and "bet" spoken by all 16 speakers in the production study (8 Ohio speakers and 8 Wisconsin speakers). The words were edited out of sentences (see 2.1.2.) and presented randomly to the listeners for identification. Twenty listeners who were born and raised in central Ohio (10 men and 10 women) and 9 listeners born and raised in southern Wisconsin (5 men and 4 women) responded to all 288 isolated words (2 vowels x 3 prosodic levels x 3 repetitions x 8 speakers x 2 dialects) in one block in one-alternative forced choice task with the choices "beet, bit, bait, bet, bat, but." After identifying the vowel, the listener then rated its goodness in terms of whether it represented a good, fair, or poor exemplar of the vowel category chosen. Sound was delivered over Sennheiser HD600 headphones to a listener seated in a sound-attenuating booth. The experiment was under computer control, using a program written in MATLAB.

4.2. Results

Only the identification data are presented here as they are most relevant to the present focus. A more exhaustive account of the perception results including vowel goodness ratings, gender-related differences, and statistical treatment can be found in Fox et al. (2006).

Overall, listeners demonstrated sensitivity to the acoustic variation in vowels as a function of differences in prosodic prominence in identifying the vowels as speakers intended. Generally, vowels produced in lower prosodic positions were misidentified more often than vowels in more prominent positions.

4.2.1. *Identification of /e/*

Figure 6 shows the average identification of /e/ by Ohio listeners responding to Ohio vowels (left panel) and to Wisconsin vowels (right panel). Of interest are both identification rates and number of confusions with neighboring vowels. For Ohio productions, listeners responded to the graded differences in acoustic vowel characteristics across prosodic positions observed in the production data. The vowels were classified as /e/ most often when they occurred in the U-position and least often when they occurred in the S-position. For Wisconsin productions, identification rates were lower. Of particular interest is that vowels in the U-position were identified as intended by the speakers only 81% of the time as compared to 93% of the time for Ohio productions. This suggests that Ohio listeners perceived a dialectal difference for vowels produced in most prominent prosodic position, which represented the clearest exemplar of the category. We will return to this in general discussion.

Results from a repeated measures ANOVA showed that these differences were significant. In particular, the significant interaction between prosodic position and speaker dialect ($F(2, 31) = 4.72$, $p = 0.020$, $\eta^2 = 0.208$) explored in subsequent post-hoc analyses demonstrated that listeners' responses to all three prosodic positions were significantly different from one another for the Ohio vowels but not for Wisconsin. For the latter, although the U- and P-positions produced significantly more /e/ responses than the S-position, they were not significantly different from one another.

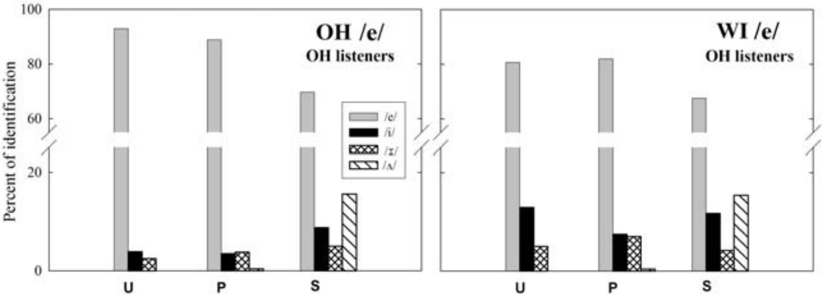


Figure 6: Identification of /e/ across three prosodic levels (U, P, S) for Ohio listeners responding to Ohio productions (left) and Wisconsin productions (right).

The results for Wisconsin listeners are shown in Figure 7. It is striking that vowels in Ohio productions were identified as intended by the speakers more often than vowels in Wisconsin productions (see the parallel findings of Labov and Ash 1997, also discussed in Labov 2001: 490). In particular, identification rate of Ohio /e/ in the U-position was much higher (97%) than that of Wisconsin /e/ (83%). Also, for Wisconsin productions, acoustic differences as a function of prosodic context did not yield differences in listeners' responses, and vowels in all three prosodic contexts were classified as intended by the speakers about 83% of the time. However, Wisconsin listeners perceived differences due to prosodic context in Ohio productions, and vowels in the lowest S-position were still perceived more often as intended by the speakers (87%) than vowels in Wisconsin productions.

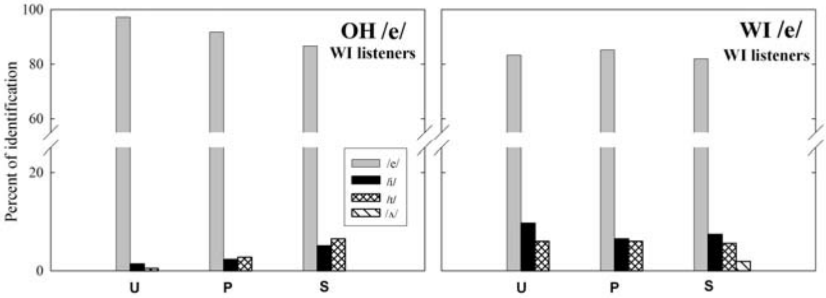


Figure 7: Identification of /e/ across three prosodic levels (U, P, S) for Wisconsin listeners responding to Ohio productions (left) and Wisconsin productions (right).

Comparing the substitution patterns shown in Figures 6 and 7, it is clear that vowels spoken by Wisconsin speakers were perceived more often as /i/ by both Ohio and Wisconsin listeners than vowels in Ohio productions. This was true particularly for vowels in the U-position. Substitutions by /ɪ/ were also prevalent, although less frequent than by /i/. A striking result was obtained for Ohio listeners' responses to the vowels in S-position. For both Ohio and Wisconsin speakers, the misidentified vowels were perceived most often as /ʌ/ and not as either /i/ or /ɪ/ as in the responses of the Wisconsin listeners. This outcome may be related to Ohio listeners' response to lesser formant movement for /e/ in the S-position as compared to the two more prominent positions. Such response most likely reflects dialectal differences in perceptual attunement to vocalic distinctions and needs to be addressed in future experiments.

4.2.2. *Identification of /ɛ/*

Displayed in Figure 8 is the average identification of the lax vowel /ɛ/ by Ohio listeners responding to Ohio and Wisconsin productions. The identification rates were very similar for both dialects. Consistently with prosodic hierarchy, vowels in the S-position were identified less often as intended by the speakers than vowels in higher prosodic positions. However, both U- and P-positions had comparable identification rates and there were no clear differences between them. The results of an ANOVA performed on the identification data showed that the effect of prosodic position was significant ($F(2, 31) = 12.15$, $p < 0.001$, $\eta^2 = 0.403$). Subsequent post hoc analyses indicated that vowels in the S-positions were identified significantly less often as /ɛ/ than vowels in either U- or P-position, which did not differ from each other.

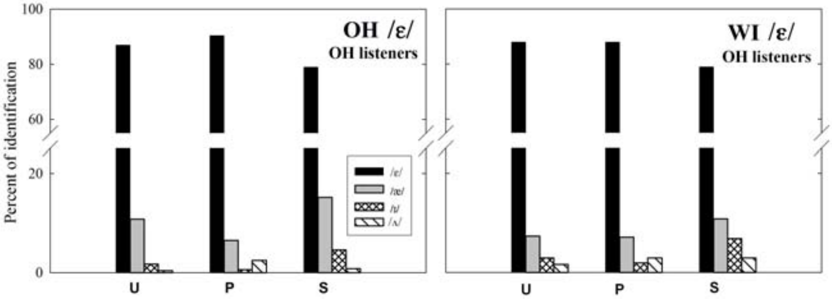


Figure 8: Identification of /ɛ/ across three prosodic levels (U, P, S) for Ohio listeners responding to Ohio productions (left) and Wisconsin productions (right).

As further shown in Figure 8, some dialectal differences were perceived. Notably, the number of substitutions by /æ/ for Ohio productions was higher than for Wisconsin productions. The substitutions for Wisconsin /e/ were more distributed across the neighboring vowels, including /æ/, /ɪ/, and /ʌ/.

Identification rates by Wisconsin listeners were slightly higher across prosodic positions and, unlike in the case of /e/, Wisconsin productions yielded a higher number of responses as intended by the speakers, particularly in the U-position. As evident in Figure 9, responses of Wisconsin listeners reflected the graded prosodic differences for Wisconsin productions but the identification pattern for Ohio vowels was similar to that for Ohio listeners. The results of an ANOVA showed that the main effect of prosodic position was significant ($F(2, 28) = 7.65$, $p = 0.002$, $\eta^2 = 0.353$), although significant differences were obtained only between the S-position and either U- or P-position, which did not differ from each other.

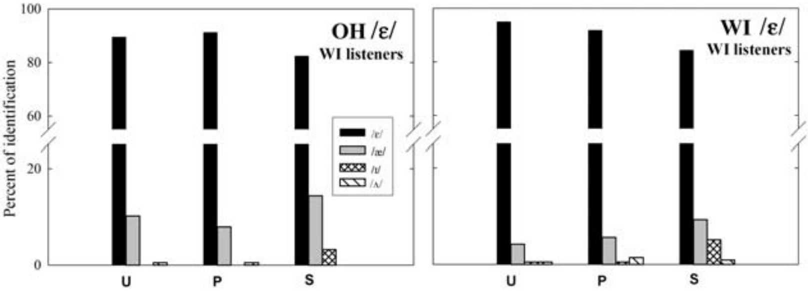


Figure 9: Identification of /ɛ/ across three prosodic levels (U, P, S) for Wisconsin listeners responding to Ohio productions (left) and Wisconsin productions (right).

As has been previously observed in the substitution pattern for Ohio listeners, the vowels spoken by Ohio speakers were confused with /æ/ more often than vowels spoken by Wisconsin speakers. For the small overall number of substitutions in classification of Wisconsin vowels, the responses were more distributed over the neighboring vowels /æ/, /ɪ/, and /ʌ/.

5. General discussion

In our exploration of possible synchronic sources of diachronic vowel changes, we examined differences in vowel characteristics as a function of variation in prosodic prominence.[4] We expected that more emphatic vowel productions would be reflected in greater vowel enhancement, which, over time, will serve as an impetus to sound change across generations of speakers from the same geographic area.

As a specific example, we selected two diachronic principles of sound change, the raising of tense vowels and possible lowering of lax vowels in chain shifts. Investigating changes to vowel characteristics as a function of differences in prosodic prominence, we found noteworthy synchronic parallels to the diachronic changes. Comparing the results from two closely related dialects of Midwestern American English, we aimed to find whether and to what extent the general trends are influenced by dialect-specific differences.

Measured acoustically, we found Ohio vowels /e/ and /ɛ/ to be longer and more fronted relative to Wisconsin vowels. We predicted that both duration differences and differences in the acoustic location may affect how the vowels are perceived by listeners born and raised in the same geographic areas as the speakers and may also affect cross-dialect vowel perception.

Measuring formant frequency changes at multiple locations equidistant in the course of the vowel's duration, we were able to detect gradual fronting and raising of the tense vowel /e/ with each position of higher prosodic prominence. This is because acoustic correlates of vowel enhancement are manifested in more extreme frequency values and, in the case of a diphthongal vowel, a greater degree of diphthongization. As could be expected, longer durations of Ohio vowels caused a greater amount of frequency change as compared to Wisconsin vowels.

An apparent contradiction with respect to raising and fronting of the vowel /e/ emerged from the data. Namely, since sound change involves both speaker and listener, the more enhanced instances of /e/, although raised and fronted in the acoustic space, should actually result in a better identification of the vowel as /e/ with each higher prosodic position. This would suggest a change in the acoustic location and no perceived vowel category change under greater amount of enhancement. Would greater prosodic prominence create a clearer exemplar of a vowel or would it cause a vowel category shift?

Perception results showed that either outcome is possible. For Ohio vowels, greater enhancement resulted in higher identification rates and lesser confusions with neighboring vowels for both Ohio and Wisconsin listeners. Moreover, longer, more fronted and more diphthongal instances of Ohio vowels were favored by Wisconsin listeners over the vowels of their own dialect, yielding higher identification rates. For Wisconsin vowels, however, more enhanced vowel variants did not provide clearer exemplars for the listeners. To the contrary, both Wisconsin and Ohio listeners perceived them more often as /i/ or /ɪ/. This suggests that Wisconsin /e/ is more likely to undergo a perceptual category shift under greater enhancement, sounding to speakers of other, non-shifting dialects more raised and fronted. This is in line with diachronic Principle I and suggests that the principles of vowel changes formulated in the past have a synchronic basis.

The question arises as to why the tense vowel /e/ tends to rise in Wisconsin English and not in Ohio. Although we can only speculate at present, the raising of /e/ may be related to the other changes undergoing in Wisconsin vowel system. While it is known that Northern Cities Shift affects short vowels only, the acoustic characteristics of long vowels have not been explored sufficiently and perception of long vowels has not been tested to rule out possible changes to long vowels in the dialect of south-central Wisconsin. Consequently, since English spoken in central Ohio is unaffected by any chain shift, we would not necessarily expect the raising of /e/ in Ohio.

Turning our attention to the second vowel investigated in this study, the lax /ɛ/, we found smaller formant frequency changes as a function of prosodic context for both Ohio and Wisconsin productions. In both dialects, the vowel is a true monophthong and we thus do not expect it to become diphthongized in more emphatic productions. Variation in prosodic prominence affected primarily vowel duration although small changes to formant frequency in more prominent positions were also observed. As a consequence of these frequency changes, the vowel was slightly lower and more fronted in more emphatic productions of both Ohio and Wisconsin speakers.

Before we discuss our perception results for /ɛ/, we need to address the issue of intelligibility scores for this vowel in other reported studies. In a study conducted in Michigan (Hillenbrand et al. 2001), listeners identified eight vowels, including /ɛ/, from a word list read by the speakers of the same Northern Cities dialect. The vowels were embedded in various consonant contexts. The intelligibility was lowest for the vowels /æ/ and /ɛ/ and the average identification rate for these vowels was about 90%. The most common misidentifications of tokens that were intended as /ɛ/ were /ɪ/ or /æ/. Comparing our results with Hillenbrand et al. (2001) data, the identification rates for /ɛ/ in the lowest prominence position were much lower, reaching 79% for Ohio listeners and 83% for Wisconsin listeners. This underscores that variation as a function of prosodic context may have great effects on vowel intelligibility and clearly supersedes variation typically found across consonant contexts.

It is important to note that the instances of vowel /ɛ/ spoken by Ohio speakers were misidentified most often as /æ/ by both Ohio and Wisconsin listeners. However, Wisconsin misidentifications of /ɛ/ were more distributed among neighboring vowel categories /æ/, /ɪ/, and /ʌ/ in the responses of both Ohio and Wisconsin listeners. This indicates that listeners were able to detect acoustic differences between Ohio and Wisconsin instances of /ɛ/. More fronted Ohio vowels yielded more confusion with /æ/ and more centralized Wisconsin vowels were more often confused with other neighboring vowels.

Diachronic Principle II, that short vowels fall in chain shifts, finds less clear synchronic support in light of the present results, although our data support the observed pattern of /ɛ/ backing currently ongoing in a number of American dialects, and therewith our revision of Principle II. Although the more emphatic productions of /ɛ/ did cause lowering of the vowel in the acoustic space, the perceptual results of this lowering in terms of a larger number of substitutions by /æ/ were somewhat greater for Ohio vowels which are assumed to be unaffected by any chain shift. Both the centralized acoustic location of Wisconsin /ɛ/ and substitutions by neighboring vowels in the perception of /ɛ/, particularly in the least prominent S-position, would suggest a gradual backing of this vowel in Wisconsin, which is consistent with changes in vowel characteristics in the Northern Cities dialects. More experimental work is needed to examine more thoroughly diachronic Principle II which, as already mentioned at the outset, seems to be ill-attested, though our results are more consistent with the backing tendency of /ɛ/ found in the Northern Cities Shift.

6. Conclusions

This study explored acoustic changes to vowels as a function of variation in prosodic prominence and examined listeners' responses to these changes with the aim of understanding how sources of systematic variation in speech can lead to dialectal vowel change over time. Generally, the results shed light on the question of how vowels change in particular directions over time. More emphatically produced vowels correlate in our data with greater vowel enhancement, which serves as a primary impetus to dialect-specific perceptual judgments about vowel quality. In the case of tense diphthongal /e/, dynamic rather than static vowel characteristics determine the nature and direction of the change. They also contribute to a lack of such change in a dialect which is not undergoing chain shifting. A more complex relationship between acoustic variation in vowel enhancement, its perception, and dialectal differences emerged for the lax monophthongal vowel /ɛ/, although the data are rather suggestive as to the direction of the change in the dialect undergoing Northern Cities Shift.

Overall, then, our results suggest a likely connection between vowel enhancement and vowel changes of the sort posited for chain shifts. We are now beginning work to pursue such connections at the seam between generations, namely in the speech of grandparents, parents, and children.

Acknowledgements

We thank Kristin Hatcher, Jennifer Delahanty and Dilara Tepeli for help with collection of the perception data. Special thanks go to Jeff Murray for editorial help. Work supported by NIH NIDCD R01 DC006871-01.

Notes

1. An anonymous reviewer reminds us that other changes do not behave in accordance with Principle II, such as a lax vowel raising pattern in New Zealand (Watson et al. 1998, Watson et al. 2000).
2. There are of course many and varied phonological approaches to prosodic prominence within sentences, the most important reviewed in Selkirk (1995). Our approach here, using three levels from utterance-initial to unstressed forms (but with non-reduced vowels) and an intermediate one to produce graded degrees of emphasis, is intended to avoid those theoretical controversies for the moment to the extent possible.
3. In fact, these stimuli are not entirely structurally parallel. In particular, as one reviewer also notes, *shark bait* forms a phonological word while *risky bets* does not. We have since revised the stimuli extensively and begun testing them. We hope to report the results soon.
4. There is a long tradition in phonetics and historical linguistics of connecting sound change to patterns of hypo- and hyper-articulation, and similar terms. In the introduction, above, and in related work (Jacewicz et al. 2006), we have suggested a reason for the apparent correlation between prosodically-driven ‘exaggeration’ and vocalic chain shifting, rooted in language learning.

Appendix

Table 1: ANOVA results for the tense vowel /e/ showing significant main effects and interactions for F1 and F2 measurements.

Effect	dF	dFe	F	p	η^2
OHIO					
Prosodic level_F1	1.3	8.0	6.80	0.025	0.531
Measurement location_F1	1.6	9.7	21.11	<0.001	0.779
Prosodic level_F1 x measurement location_F1	2.7	16.3	5.86	0.008	0.494
Prosodic level_F2	1.4	8.5	28.40	<0.001	0.826
Measurement location_F2	1.3	7.5	43.60	<0.001	0.879
Prosodic level_F2 x measurement location_F2	2.1	12.9	6.460	0.010	0.519
WISCONSIN					
Prosodic level_F1	1.6	9.4	13.86	0.002	0.698
Measurement location_F1	1.7	10.1	18.05	0.001	0.751
Prosodic level_F1 x measurement location_F1	ns				
Prosodic level_F2	1.7	10.3	55.35	<0.001	0.902
Measurement location_F2	2.4	14.6	72.18	<0.001	0.923
Prosodic level_F2 x measurement location_F2	2.4	14.3	4.95	0.019	0.452

Table 2: ANOVA results for the lax vowel /ɛ/ showing significant main effects and interactions for F1 and F2 measurements.

Effect	dF	dFe	F	p	η^2
OHIO					
Prosodic level_F1	1.9	11.4	5.91	0.016	0.496
Measurement location_F1	2.2	13.0	8.64	0.004	0.590
Prosodic level_F1 x measurement location_F1	ns				
Prosodic level_F2	1.7	10.4	9.78	0.003	0.620
Measurement location_F2	ns				
Prosodic level_F2 x measurement location_F2	ns				
WISCONSIN					
Prosodic level_F1	1.6	9.4	5.04	0.038	0.457
Measurement location_F1	2.1	12.7	44.88	<0.001	0.882
Prosodic level_F1 x measurement location_F1	ns				
Prosodic level_F2	1.8	10.6	9.20	0.006	0.605
Measurement location_F2	ns				
Prosodic level_F2 x measurement location_F2	2.2	12.4	6.47	0.010	0.519

References

Cho, Taehong and Patricia A. Keating
2001 Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics* 29: 155-190.

Ferguson, Sarah Hargus and Diane Kewley-Port
2002 Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-imparted listeners. *Journal of the Acoustical Society of America* 112: 259-271.

Fox, Robert Allen, Ewa Jacewicz, and Joseph Salmons
2006 Prosodically induced phonetic variations in vowels: A source of language change? In *The Bill Question*, Howard I. Aronson, Donald L. Dyer, Victor A. Friedman, Daniela S. Hristova, and Jerrold M. Sadock (eds.), 87-110. Bloomington, Ind.: Slavica.

Gordon, Matthew
2001 *Small-Town Values and Big-City Vowels: A Study of the Northern Cities Shift in Michigan.* (*Publications of the American Dialect Society*, 84) Durham: Duke University Press.

Gordon, Matthew
2002 Investigating chain shifts and mergers. In *The Handbook of Language Variation and Change*, J.K. Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), 244-266. Maldon, Mass. and Oxford: Blackwell.

Hillenbrand, James M., Laura A. Getty, Michael J. Clark, and Kimberlee Wheeler
1995 Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97: 3099-3111.

Hillenbrand, James M., Michael J. Clark, and Terrance M. Neary
2001 Effects of consonant environment on vowel formant patterns. *Journal of the Acoustical Society of America* 109: 748-763.

Jacewicz, Ewa, Joseph Salmons, and Robert Allen Fox
2006 Prosodic prominence effects on vowels in chain shifts. *Language Variation and Change* 18: 285-316.

Jacewicz, Ewa, Robert Allen Fox, and Joseph Salmons
2007 Vowel duration in three American English dialects. *American Speech* 82: 367-385.

Labov, William
1994 *Principles of Linguistic Change, 1: Internal Factors.* Oxford: Blackwell.

Labov, William
2001 *Principles of Linguistic Change, 2: Social Factors.* Oxford: Blackwell.

Labov, William and Sherry Ash
1997 Understanding Birmingham. In *Language Variety in the South Revisited*, Cynthia Bernstein, Thomas Nunnally, and Robin Sabino (eds.), 508-573. Tuscaloosa: University of Alabama Press.

Labov, William, Sherry Ash, and Charles Boberg
2006 *The Atlas of North American English: Phonetics, Phonology and Sound Change.* Berlin and New York: Mouton de Gruyter.

Labov, William, Malcah Yeager, and Richard Steiner
1972 *A Quantitative Study of Sound Changes in Progress.* Philadelphia: U.S. Regional Survey.

Milenkovic, Paul
2003 TF32 software program. University of Wisconsin, Madison.

Onaka, Akiko
2003 Domain-initial strengthening in Japanese: and acoustic and articulatory study. In *Proceedings of the 15th International Congress of Phonetic Sciences*, Maria-Josep Solé, Daniel Recasens, and Joaquín Romero (eds.), 2091-2094. Barcelona, Spain.

Selkirk, Elizabeth
1995 Sentence prosody: Intonation, stress, and phrasing. In *The Handbook of Phonological Theory*, John A. Goldsmith (ed.), 550-569. Oxford: Blackwell.

Sievers, Eduard
1876 *Grundzüge der Lautphysiologie: Zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen.* 2nd ed. *Grundzüge der Phonetik.* Leipzig: Breitkopf and Härtel. Second edition 1881.

Stockwell, Robert P.
1978 Perseverance in the English vowel shift. In *Recent Developments in Historical Phonology*, Jacek Fisiak (ed.), 337-348. The Hague: Mouton.

Watson, Catherine I., Jonathan Harrington, and Zoe Evans
1998 An acoustic comparison between New Zealand and Australian English vowels. *Australian Journal of Linguistics* 18: 185-207.

Watson, Catherine I., Margaret Maclagan, and Jonathan Harrington
2000 Acoustic evidence for vowel change in New Zealand English. *Language Variation and Change* 12: 51-68.

Variable quality of the Czech lateral liquid: A perception experiment with young Czech listeners

Šárka Šimáčková

1. Introduction

People's speech patterns shift and with them shift their pronunciation norms, in other words their notion of what sounds good, correct, or standard and what does not. A prescriptive linguist, and sometimes a concerned member of lay public, keen on maintaining the normative pronunciation, concedes that a norm began to fluctuate and declares the need to reverse this corruptive process.

The paper describes an instance of norm fluctuation that is tolerated by speakers but rejected by language professionals who work with *orthoepic* rules – commonly accepted rules of standard pronunciation (Hála 1967, also e.g. Palková 1994a, Lechta 1995, Vyštejn 1995). The norm in question concerns the pronunciation of the lateral liquid in Czech. At the beginning was a casual observation that many, though by no means all, first year students whom I encounter in my courses pronounce dark *l* in Czech. An informal query among fellow phoneticians revealed that others have noticed the darkening of *l* in casual speech before me. It also transpired that the strong orthoepic sentiment present in Czech phonetics cannot be reconciled with such apparently sloppy and uneducated pronunciation. This made me wonder whether the non-expert population would have equally strong feelings about this phenomenon.

1.1. Descriptions of Czech *l*

Synchronically, the system of liquids in Czech is simple, contrasting the lateral with the trilled *r*.[1] The old contrast between the palatalized and velarized *l* is preserved to some extent in small dialectal areas in the east of the country, in Silesia and East Moravia (Bělič 1972). Geography does not seem to be relevant to the occurrence of the dark *l* in the speech of young Czechs though Common Czech as spoken in Central Bohemia is a good place to start looking for such weakening. This variety of Czech is known for more lax, open pronunciations.

Phonetically, the Czech lateral is described as an apico-alveolar lateral approximant of light or neutral quality (Hála 1962, Dankovičová 1999). Palková (1994b) describes it as an apico-alveolar lateral fricative. Apart from the syllabic *l* no allophonic variation is ever discussed. When non-dialectal dark realizations of *l* are mentioned in the phonetic and phonological literature, they are treated as a speech pathology, representing a form of *lambdacism.*

1.2. Lambdacism

Lambdacism, defined as defective pronunciation of the sound [l] (Dvořák 2001), covers a range of pronunciations. Some authors mean by lambdacisms vocalization of *l* but the term may also refer to deletion of *l* or its substitution with /j/, /h/, or /v/ (Vyštejn 1995). The latter substitutions are more common in small children. Adults are more likely to produce vocalic sounds with *o*-like or *u*-like resonance. This is probably what Sovák means when he calls the dark *l* persistent ["tvrdošíjné"] (Sovák 1978: 175).

No statistics describing speech defects in adults are available but the number of children diagnosed with lambdacisms after they started school appears to be rising. In his *Pronunciation of Standard Czech* from 1967, Hála does not include lambdacism in the list of consonantal defects at all and only mentions the hard *l* as a regional feature limited to Northeast Moravia. The defective *l* does not appear in the summary of 1960 study of 3389 elementary school children by Kopecký, Cop and Nosek (in Sovák 1978). According to Ohnesorg (1956) lambdacism represents 2.6% of all speech defects. In the 1989 study reported by Benešová and colleagues in the speech therapy yearbook of 1991 5% out of the 5599 tested first-graders were diagnosed with lambdacism. Compared to the rhotics and sibilants the percentage is still relatively small (See the top line in Table 1). The outcome of a similar survey carried out a decade later with 6442 children (Půstková 2000) shows that the number of children with a poor command of articulation of *l* had more than doubled while other defects had increased less noticeably or had not increased at all.

Our attention is drawn to this disproportionate rise in defective *l* by Volín who is worried by the "increasing number of sloppy *l*'s in young people" [narůstající množství těch „nepořádných“ [l] u mladých lidí] (Volín 2002: 7). Apparently the country has been hit with an epidemic of lambdacism. Or

Table 1: Percentages of first-grades diagnosed with defective pronunciation in 1989 and 1999 (Benešová et al. 1991, Půstková 2000).

	l	r	ř	ts s z	tʃ ʃ ʒ	t d n
1991	5	13	17	16	9	7
2000	11	23		17	11	8

maybe not. I suggest that the growing number of school children diagnosed with a defective *l* can be viewed as reflecting a more general trend towards a variable pronunciation of the Czech lateral. More and more younger speakers allow a lenitive ‘velarization” process that many older or more conservative speakers effectively inhibited. Speech patterns are shifting while the prescribed norms remain unchanged.

Compared to the light *l* articulated with the tongue tip against the alveolar ridge, the lenited dark *l* involves a diminished or completely lost apico-alveolar contact and a retraction of the tongue body, with a possible raising of the back of the tongue towards the velum (Giles and Moll 1975). The inverted commas around velarization are a reminder that the upward movement of the tongue dorsum towards the velum may not be necessary in the articulation of the dark variant (Sproat and Fujimura 1993). Complete elimination of the coronal gesture results in a vowel-like segment. Acoustically, the weakening of *l* is manifested in a much lower F2, a somewhat higher F1 and the decreased distance between F1 and F2 (Espy-Wilson 1992). The corresponding auditory impression is that of a darker sound.

There is nothing unusual about an alveolar *l* developing a dark retracted variant. The propensity of both liquids, lateral and rhotic, to vary between light and dark quality has been documented cross-linguistically – /l/ e.g. in English (Cruttenden and Gimson 2001), Dutch (Booij 1995), Portuguese (Mateus and Pardal 2000), Latin (Schein, B. and D. Steriade 1986); /r/ e.g. in German dialects (Howel 1987). When the light and dark liquid alternate synchronically, the dark variant is preferred post-vocalically while the light variant occurs pre-vocalically.

Dark qualities of both *l* and *r* can occur in Czech but they seem to fall outside the boundaries of standard pronunciation. The situation is quite clear in the case of the rhotic. Substituting the standard apico-alveolar trill with a velar approximant and a velar or uvular trill is regarded as a speech pathology – *rhotacism* – by Czech speakers of different ages.[2] These substitutions are readily recognized, often mimicked and even stigmatized. They have earned a popular name (*ráčkování* derived from the word *rak*, 'crayfish'). The same cannot be claimed about Czech speakers' judgments of the dark *l*. It is not certain that there is a stereotypical evaluation of such pronunciation. Thus in this paper I ask the following: How tolerant of dark *l* are young adult speakers who are exposed to the variable quality of *l* in speech of their peers and who may or may not produce it themselves? Or do they perceive it as normal, sloppy, or defective?

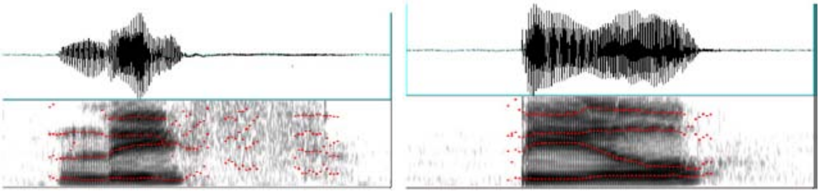
Since Czech listeners' judgments of the pronunciation of *r* are categorical and predictable they were used in the perception study reported here as a baseline to which Czech listeners' responses to the dark *l* were compared.

2. An earlier production study

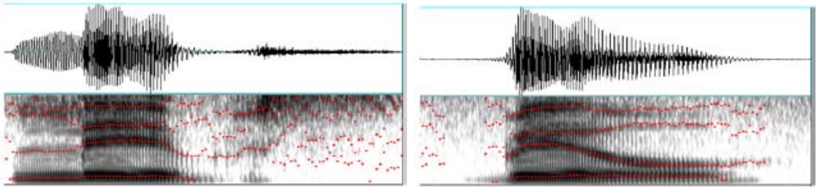
An earlier study carried out with 25 female college students supported my casual observations of the dark quality of the lateral sound. Both light and dark variants of the phoneme were observed in the data, their distribution more or less conforming to the pattern found crosslinguistically: light *l* in the onset and dark *l* in the coda position. Figure 1 shows that, irrespective of the orthoepic norm, the post-vocalic Czech *l* can have similar acoustical attributes as dark *l* in English (low F2, small F2-F1 difference). Spectrographs of British English words *bill, lives* and Czech *pil,* 'he drank', *lis,* 'a press' are compared. Both initial and final *l*'s are similar in the two languages (one possible difference could be in the duration of the segment, with the English post-vocalic *l* being longer in this particular instance).

However, the study also found occurrence of some non-light laterals in syllable onsets and a great variability in the pronunciation of *l* in the codas. Dark *l* occurred in the sample of post-vocalic *l*'s as often as it did not. There was variation among speakers, who are of the same gender and similar age, as well as variation in speech of individuals. Table 2 gives the number of speakers preferring each variant. What these data show is that actual pronunciation of the lateral phoneme varies more than what textbook descriptions of standard Czech would have us believe.

The variability in pronunciation of *l* contrasts with the relative uniformity in Czech speakers' pronunciation of the rhotic liquid – trilled *r*. In two studies (Šimáčková 2002, 2003) I examined trills in speech of two radio presenters and in speech elicited from three females of different ages. These speakers realized 90-95% or their *r*'s as a single tap trill. In sum, pronunciation of the rhotic liquid shows little variation and there is a clear sense of the dark *r* being substandard. Pronunciation of the lateral phoneme on the other hand seems to be rather unstable. The question asked here is whether the production fluctuations will be paralleled by a less definite rejection of the dark *l*.



(a) Czech *lis* (b) Czech *pil*



(c) English *lives* (d) English *bill*

Figure 1: Pre and post-vocalic *l*'s in Czech *lis*, *pil* and English *lives*, *bill*.

Table 2: Realization of final *l* by young adults.

realization of /l/	Number of speakers	%
[l]	10	40%
[ɫ] [o̯]	9	36%
[l] [l̵] [o̯]	6	24%

3. The current perception study

Below I report results of a new study that looks at the lateral liquid in perception. I wanted to find out how sensitive young Czechs are to the more or less dark quality of the lateral phoneme in the speech of their peers. I took as my starting point one of the basic tenets of Natural Phonology (Stampe 1979), according to which phonological processes operate in perception as well as in production, undoing their own effects and enabling the listener to perceive sounds as intended by the speaker. The listener automatically factors out all phonetically motivated substitutions. If the lenition of *l* is an active process in a Czech listener's phonology, the listener should not be aware of the dark quality of the post-vocalic *l*'s. If, however, the listener has inhibited the lenition process in the course of learning her language, she will notice dark *l*'s in the same way that she notices non-native substitutions in the accented speech of foreigners.

I asked a group of young Czechs to listen and judge the quality of the lateral in a set of words. I wanted to know (1) whether they would notice dark *l*'s in speech of their peers at all; (2) if so, whether they would be more likely to notice dark *l*'s when they occur outside the appropriate phonological context, i.e. pre-vocalically, or (3) whether they would notice dark *l*'s even in the post-vocalic position that is most susceptible to weakening. I assumed that if they did indeed notice the dark quality, it would be because it did not match their pronunciation norm and they would judge such *l*'s as poor examples. In order to get a handle on listeners' judgments I compared their perception of the lateral liquid to their perception of the rhotic liquid.

3.1. Participants, materials, task

Thirty listeners, all college students aged between 19 and 27, participated in the study. In order to get 30 subjects, the original restriction on gender was ignored and seven men were included.

The stimuli consisted of isolated words that were recorded into a computer by 8 talkers, 4 women and 4 men, aged 19-25. The words contained an *l* or an *r* in the word initial, medial, and final positions in the context of vowels /i/ and /a/. The combination of position and vowel context yielded 6 lateral and 6 rhotic test words listed in Table 3.

Table 4 summarizes information about the lateral stimuli. The phonetic symbols, believed to be the closest representatives of the tokens, were chosen on the bases of auditory impression, acoustic inspection, and observation of the talkers' mouths during articulation.

Four talkers recorded the lateral set. Two have so-called "normal" pronunciation but they both produce dark *l* in word-finally. Standard Talker 1 produces initial and medial *l*'s of a middle, non-light and non-dark quality, except for light *l*'s before /i/. Her final *l*'s are dark but consonantal. Standard Talker 2 has middle-quality *l*'s initially and medially with the exception of a distinctly light *l* in *lis*. His final *l*'s are dark, fully vocalized. The remaining two talkers were both diagnosed with lambdacisms sometime during elementary school. The female Non-standard Talker 1 substitutes a more or less round back vowel for *l* in all positions except initially and before a high front vowel where she pronounces a consonantal sound of dark quality [ɫ]. The male Non-standard Talker 2 prefers a rounded vowel in *bál* and *pil* and a dark consonantal sound in *lát, lis, sálá,* and *myli*. Observing his mouth during pronunciation suggested that the latter was lamino-dental.

Table 3: Stimuli. The high front vowel /i/ is represented by the letters "i" and "y" in Czech. A stroke above a vowel indicates length.

context	lateral stimuli	context	rhotic stimuli
#la	*lát*, 'to scold'	#ra	*rád*, 'glad'
#li	*lis*, 'a press'	#ri	*rýč*, 'a spade'
a.la	*sálá*, 'it glows'	a.ra	*čára*, 'a line'
i.li	*myli*, 'they washed'	i.ri	*sýry*, 'cheeses'
al#	*bál*, 'a dancing ball'	ar#	*bar*, 'a bar'
il#	*pil*, 'he drank'	ir#	*sýr*, 'a cheese'

Table 4: Quality of the lateral in the stimuli from each talker.

Standard		Non-Standard	
Talker 1 female	Talker 2 male	Talker 1 female	Talker 2 male
l ɫ	l o̯	ɫ ʊ̯	ɫ̪ ʊ̯

Two standard talkers recorded the rhotic set. They differ from each other in that Standard Talker 1, a man, only produces a single contact trill while Standard Talker 2, a woman, has also multiple contact trills, especially in the strong initial position. The speech of Non-standard Talkers 1 and 2 was undoubtedly marked by rhotacism. Both of them reported visits to a speech therapist in their childhood and both produced dark retracted *r*'s, varying between uvulo-velar trills and velar approximants.

To summarize, four groups of stimuli were created, including two speakers each: non-standard rhotic, standard rhotic, non-standard lateral, and standard lateral.

Peritoneum 1.0 software for perception experiments (Duběda and Votrubec 2004), was used to present the 48 stimuli. The listeners heard the words over earphones. They controlled the speed of presentation by clicking on an icon on the screen to hear the current stimulus and then clicked on the evaluation scale. A new stimulus occurred only after the evaluation of the preceding one was completed. No limit was set on how many times a stimulus could be played. The 48 words were presented three times in random order, i.e. 144 judgments were made by each listener. She or he evaluated pronunciation of each *l* and *r* on a five-point scale, going from "1" for good pronunciation, "2" for fairly good, "3" for tolerable, to "4" for fairly poor and "5" for poor pronunciation.

3.2. Results

For each listener the evaluation points for a particular stimulus were added up. A stimulus consistently evaluated by a listener as good had 3 points (1+1+1) while a stimulus consistently evaluated as poor had 15 points (5+5+5). The added up responses from our 30 listeners to the 48 stimuli were submitted to a cluster analysis. As is shown in Figure 2, five clusters emerged. They are numbered from 5 to 1 parallel to the reversed evaluation scale, Cluster 5 containing stimuli generally perceived as poor, and Cluster 1 stimuli generally perceived as good. The tokens showing at the bottom of the figure are coded as follows: shorthand for each liquid in a particular position and vowel context (e.g. "ri" = the initial *r* in the word *rýč*) is followed by speech quality (s = standard, n = non-standard), and the talker number (1 or 2).

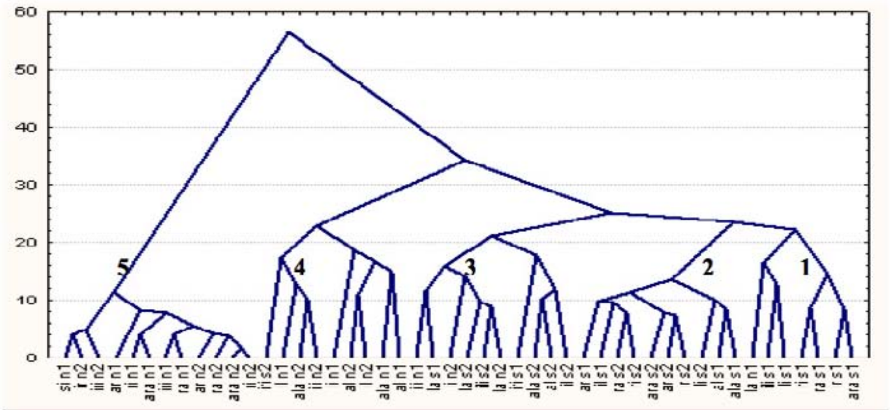


Figure 2: Results of cluster analysis.

Distribution of the stimuli across the five clusters and eight speakers is tabulated in Table 5. Gray shading emphasizes that no tokens fell into the particular cluster.

Table 5: Distribution of rhotic stimuli across clusters and talkers.

	Non-standard						Standard					
	Talker 1			Talker 2			Talker 1			Talker 2		
	INI	MID	FNI	INI	MID	FIN	INI	MID	FNI	INI	MID	FIN
C5	ra	a.ra	ar	ra	a.ra	ar						
	ri	i.ri	ir	ri	i.ri	ir						
C4												
											i.ri	
C3												
								i.ri				
C2									ar	ra	a.ra	ar
										ri		ir
C1							ra	a.ra				
							ri		ir			

Table 6: Distribution of lateral stimuli across clusters and talkers.

	Non-standard						Standard					
	Talker 1			Talker 2			Talker 1			Talker 2		
	INI	MID	FNI	INI	MID	FIN	INI	MID	FNI	INI	MID	FIN
C5												
C4		a.la	al		a.la	al						
	li		il		i.li	il						
C3				la			la			la	a.la	al
		i.li		li							i.li	il
C2								a.la	al			
									il	li		
C1	la											
							li	i.li				

The Chi-square procedure was used to evaluate the data. Table 7 gives results of four Chi-square tests that compared listeners' responses to the four groups of stimuli, each group totaling 1080 judgments (6 words x 2 talkers x 3 presentations x 30 listeners).

Table 7: Chi-square tests comparing the four main groups of stimuli.

	Camparison	df	χ^2	p value
1	non-standard *r* standard *r*	4	1614.03	≤ 0.001
2	non-standard *r* non-standard *l*	4	1462.69	≤ 0.001
3	non-standard *l* standard *l*	4	105.81	≤ 0.001
4	standard *r* standard l	4	81.90	≤ 0.001

The group of words that is clearly set off from the rest is the leftmost Cluster 5.

It contains the stimuli with the highest scores (13.5-14.6), that means with the worst ratings. The twelve stimuli are exclusively words produced by the two speakers with rhotacism. The Chi-square test comparing listeners' responses to *r* of the standard talkers with their responses to *r* of the non-standard talkers yielded a significant result (see Comparison 1 in Table 7). Fulfilling the expectation, the listeners proved to have a strict norm for the rhotic liquid.

As is evident from Table 5 talkers with the tongue tip rhotic were perceived quite differently. Their *r*'s were judged as fairly good or good. The only exception is the intervocalic *r* between high vowels in the word *sýry* which was regarded by many merely as tolerable (Talker 1) or even as fairly poor (Talker 2). A possible explanation might be the presence of some frication during the brief contact. All other *r*'s from Standard Talker 2 who produced multiple contact trill fell into Cluster 2. Some listeners reacted negatively to her hyperarticulated *r*'s and judged them as one step worse than the single tap *r*'s of Standard Talker 1. Four out of the six *r*-stimuli from this talker were evaluated at the top end of the scale. To the listeners they represented good pronunciation. In other words the single-tap *r* is these listeners' norm for the rhotic liquid. Our listeners' responses to *r* are not surprising. On the contrary, they reflect a typical reaction of a Czech speaker to standard apical versus non-standard retracted *r*. As such they can provide us with a certain perspective in discussing the listeners' judgments of *l*.

The first obvious difference between the lateral versus rhotic stimuli involves evaluation of the talkers with a speech defect. The listeners could pick out the talkers with the defective dark *l* but did not judge this deviation from the norm as strictly as the defective dark *r*. None of the laterals produced by the two talkers with lambdacisms appears in Cluster 5 although two thirds of the stimuli – eight of their *l*'s – are in Cluster 4. Three stimuli fell into Cluster 3 (*li, li, i.li*) and one even into Cluster 1 (*la*). This, however, does not mean that the listeners were more tolerant towards dark *l*'s in the onset, rather they picked up on the lighter quality of the lateral in these tokens. The talkers were simply able to produce *l*'s that are closer to the norm in the favorable positions. The relative lenience of judgments of the non-standard dark *l* compared to the non-standard dark *r* was confirmed statistically. A Chi-square test performed on listeners' responses to the talkers with lambdacisms and their responses to the talkers with rhotacism yielded a significant result (see Comparison 2 in Table 7).

The second difference in our listeners' responses to the two liquids is evident in the less categorical split between standard and non-standard talkers in the lateral set. While in the rhotic set only two out of the twelve standard stimuli elicited responses falling outside Clusters 1 and 2, half of the lateral stimuli from the standard talkers sank into Cluster 3. Five words come from the Standard Talker 2; only his distinctly light *l* in *lis* was judged as better. The sixth word is the *l* in *lát* from the Standard Talker 1. The rest of her stimuli show a two way split. The final *l*'s and *l*'s between low vowels fell into Cluster 2 and the initial *l*'s and *l* between high vowels into Cluster 1. Again, the difference between listeners' evaluations of the standard *r* and standard *l* is significant (see Comparison 4 in Table 7). Even so, the laterals produced by the talkers with lambdacisms were perceived as worse than the laterals from the standard talkers (see Comparison 3 in Table 7).

3.3. Comments

Two observations suggest a positive answer to the initial question whether young speakers of Czech notice the dark *l*: (1) Non-standard stimuli received significantly worse ratings – our listeners recognized the talkers with lambdacisms and (2) they judged *l* from the standard talkers, who allow post-vocalic weakening, as worse than the standard *r*.

Further details in the data indicate that listeners were more likely to notice the dark *l* when it was pre-vocalic. The initial *l* in *lis* was pronounced as dark but consonantal by the talkers with lambdacisms. The standard talkers produced a distinctly light variant (undoubtedly due to coarticulation with /i/). According to Table 6 the standard pronunciation received much better judgments. This is corroborated by a Chi-square analysis (comparing 180 responses to each group of talkers: 1 stimulus x 3 presentations x 2 talkers x 30 listeners; χ^2=24.16, df = 4, p $\leq$ 0.001). In addition, the middle quality *l* produced by the standard talkers before the low vowel in *lát* was also judged as worse compared to their light *l* in *lis* (χ^2=14.31, df = 4, p $\leq$ 0.01).

There is some indication that the post-vocalic dark *l* went unnoticed by the listeners. We can compare standard talkers' dark *l* in *pil* with their light *l* in *lis*. For each talker the post-vocalic *l* falls one cluster below the pre-vocalic one but when we compare listeners' responses to *lis* versus *pil* statistically, the result is not significant (χ^2=7.05, df = 4, p $\geq$ 0.05). On the other hand, the consonantal dark *l* of Standard Talker 1 in *pil* and *bál* obtained significantly better ratings than the vocalized *l* of Standard Talker 2 (χ^2=63.72, df = 4, p $\leq$ 0.001). This suggests that the listeners were sensitive to the degree of weakening of the lateral in the final position.

4. Conclusion

Based on the understanding of phonological processes by Natural Phonology I reasoned that if dark *l*'s go unnoticed in perception it is because the process that substitutes them for the lateral phoneme is active in a listener's phonology. I further reasoned that if listeners do not notice a substitution, they will rate a given pronunciation as good. Finally, I assumed that the lenition of *l* in speech of young Czechs would follow universal tendencies. In that case the post-vocalic dark *l* would be rated as better than the pre-vocalic dark *l* because it occurred in the appropriate context. Above we saw that the listeners were sensitive to dark *l*'s but did not judge them as strictly as dark *r*'s. The light [l] was preferred pre-vocalically. In the post-vocalic position, the dark consonantal [ɫ] was generally perceived as fairly good while the vocalized realization [o̯] was judged as worse but still tolerable. All in all the young listeners do not seem to share the negative sentiment of language professionals.

The post-vocalic weakening of /l/ reported here for young Czech speakers may not be a completely new development. Some gradient phonetic effect, albeit imperceptible, might be expected in speakers across generations (though without appropriate data this remains a mere speculation). In younger speakers the syllable-final position yields a more categorical result. When and how such substitution started spreading remains an open question. Comments from phoneticians and speech pathologists on the phenomenon appeared within the last decade and they point primarily to youngsters as the main culprits. To this we may add that for young speakers pronunciation of /l/ is in the state of flux. Variable pronunciation of postvocalic /l/ reported earlier for young adults is complemented here by judgments of such pronunciations as neither categorically poor nor good.

Notes

1. The third possible member is *ř* also classified as a liquid. However, the sound behaves phonologically as an obstruent and thus the description "trilled fricative" seems to be more accurate.
2. It is important to keep in mind that the deviation from the articulatory norm does not always involve tongue retraction. Table 1 does not give details of the type of rhotacisms and lambdacisms noticed.

References

Benešová, E., Bubeníčková, M., Janotová N., and P. Janota
1991 Průzkum výslovnosti u školních začátečníků v Praze v roce 1989. In *Česká Logopedie* 1991, 50-53. Praha: Knižní podni-katelský klub.

Bělič, Jaroslav
1972 *Nástin České Dialektologie*. Praha: Státní pedagogické nakla-datelství.

Booij, Geert
1995 *The Phonology of Dutch*. Oxford: Oxford University Press.

Dankovičová, Jana
1999 Czech. In *Handbook of the International Phonetic Association*, 70-73. Cambridge: Cambridge University Press.

Duběda, Tomáš, and Jan Votrubec
2004 Configurable software for individual perception testing. In *Czech-German Workshop*, 45-49. Praha: Akademie Věd České Republiky.

Dvořák, Josef
2001 *Logopedický Slovník*. Žd'ár nad Sázavou: Logopedické centrum.

Espy-Wilson, Carol Y.
1992 Acoustic measurements for linguistic features distinguishing the

semivowels /w,j,r,l/ in American English. *Journal of Acoustical Society of America* 92(2): 736-757.

Giles, S. B., and Moll, K. L.
1975 Cinefluorographis study of selected allophones of English /l/. *Phonetica* 31(3-4): 206-227.

Cruttenden, Alan, and A. C. Gimson
2001 *Gimson's Pronunciation of English*. 6th ed. London: Arnold.

Hála, Bohuslav
1962 *Uvedení do Fonetiky Češtiny na Obecně Fonetickém Základě*. Praha: Nakladatelství Československé Akademie Věd.

Hála, Bohuslav
1967 *Výslovnost spisovné češtiny I*. Praha: Academia.

Howel, Robert. B.
1987 Tracing the origin of uvular R in the Germanic languages. *Folia Linguistica Historica* 7: 317-349.

Lechta, Viktor
1995 *Diagnostika Narušenej Komunikačnej Schopnosti*. Osveta: Martin.

Mateus, Maria Helena Mira, and Ernesto D'Andrade Pardal
2000 *The Phonology of Portuguese*. Oxford: Oxford University Press.

Ohnesorg, Karel
1956 Frekvence dyslalií a jejich fonetický výklad. In *Česko-slovenská Logopedie*, 100-113. Praha: Státní Pedago-gické Nakladatelství.

Palková, Zdena
1994a *Stručná Pravidla České Ortoepie*. Praha: Český rozhlas – odbor vzdělávání.

Palková, Zdena
1994b *Fonetika a Fonologie Češtiny*. Praha: Univerzita Karlova.

Půstková, Zuzana
2000 Zpráva o průzkumu správné výslovnosti u žáků prvních ročníků ZŠ v roce 1999/2000. *Česká Logopedie* 2000, 100-111. Praha: Galén.

Schein, Barry, and Donca Steirade
1986 On geminates. *Linguistic Inquiry*17: 691-744.

Sovák, Miloš
1978 *Logopedie*. Praha: Státní Pedagogické Nakladatelství.

Sproat, Richard, and Osamu Fujimura
1993 Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21: 291-311.

Stampe, David
1979 A Dissertation on Natural Phonology. New York: Garland.

Šimáčková, Šárka
2002 Retracted pronunciation of the Czech alveolar trill: Why phonologists need phonetic labs. In *New Challenges for Natural Linguistic*s,

Katarzyna Dziublaska-Kolaczyk, and Jacek Weckwerth (eds.), 233-249. Munich: Lincolm Studies in Theoretical Linguistics 30.

Šimáčková, Šárka
2003 Variability in the realization of Czech trilled *r*. In *Investigations into Formal Slavic Linguistics. Proceedings from Fourth European Conference on Formal Description of Slavic Languages*, Kosta Peter, Blaszczak, Joanna, Frasek Jens, Geist Ljudmila and Maryena Zygis (eds.), 119-135. Frankfurt am Main: Peter Lang.

Volín, Jan
2002 Čtyři scénáře vývoje české laterály. *Čeština Doma a ve Světě* 1: 7-13.

Vyštejn, Jan
1995 *Vady Výslovnosti*. Praha: Státní Pedagogické Nakladatelství.

Patterns of lenition in Brazilian Portuguese[1]

Thaïs Cristófaro-Silva & Daniela Oliveira Guimarães

1. Introduction

A challenge imposed on theories of sound variation and change is to identify the range of properties that determine the transition from one phonetic state to another one (Labov 1994: 543). A great number of current research in sound change has been concerned with identifying patterns of sound change and their source (Browman and Goldstein 1992, Docherty et al. 1997, Zsiga 1995). This paper aims to contribute towards this line of research which has recently been evaluating conventional categories of description employed in phonetics and phonology. We will consider cases of lenition in Brazilian Portuguese (henceforth BP). We demonstrate that fine phonetic detail is crucial to determine the development of sound variation and change which may yield the emergence of a new phonetic category.

We also discuss the role of token frequency effects in the implementation of sound changes. We will examine the claim that regular sound changes progress more quickly in items of high token frequency (Bybee 2001: 11). We will argue that patterns of lexical diffusion are important to understand fully how lenition of affricates is implemented in BP.

In the first section we present general information about palatalization and related phenomena in BP. Section 2 presents an acoustic analysis which examines continuous phonetic properties that may be captured in spectrograms. It is argued that the observed patterns reveal important information about the transition from one phonetic state to another one. It is then suggested that phonetic detail is inherent to variation and must be taken into account if one is to fully understand sound changes. In section 3 lenition is examined in relation to token frequency and stress distribution. We argue that patterns of lexical diffusion are implemented through a combination of structural information and token frequency rates.

An important contribution of this paper is to recognize the emergent character of phonetic categories yielding the transition from one phonetic state to another one. The results presented also support the view that sound change is implemented in a phonetically gradual fashion through patterns of lexical diffusion (Bybee 2000, 2001).

2. Palatalization and related phenomena in BP

A diachronic analysis of Brazilian Portuguese shows that affricates were not reported until recently in any variety of Portuguese, except in some loan words (Teyssier 1987). Consider Table 1.

Table 1: Palatalizing and non-palatalizing dialects in BP.

		Palatalizing varieties	non-palatalizing varieties	orthography	glossa
	a	ˈtʃipʊ	ˈtipʊ	*tipo*	‘type’
	b	ˈtʃĩtə	ˈtĩtə	*tinta*	‘paint’
tʃ	c	ˈahtʃɪ	ˈahtɪ	*arte*	‘art’
	d	ˈpatʃɪ̯u	ˈpatɪ̯u	*pátio*	‘yard’
	e	tʃiˈatɾʊ	tiˈatɾʊ	*teatro*	‘theatre’
	f	ˈdʒitʊ	ˈditʊ	*dito*	‘said’
	g	ˈdʒĩdə	ˈdĩdə	*dinda*	‘woman’s name’
dʒ	h	ˈaɦdʒɪ	ˈaɦdɪ	*arde*	‘sting’
	i	ˈĩdʒɪ̯u	ˈĩdɪ̯u	*índio*	‘Indian’
	j	dʒiˈbaɪ̯ʃʊ	diˈbaɪ̯ʃʊ	*debaixo*	‘under’

Palatalization is currently the major distinctive feature of Portuguese spoken in Brazil by which alveolar stops occur systematically as affricates when followed by a high front vowel as illustrated in Table 1. The high front vowel may be stressed (1 a, b, f, g), unstressed pretonic (1 e, j) or unstressed postonic (1 c, d, h, i), oral (1 a, c, d, e, f, h, i, j), nasal (1 b, g) or a glide (1 d, i), as it is shown in Table 1.

Palatalization of alveolar stops is seen as an innovative feature of BP in relation to European Portuguese and other African and Asian varieties, although there has been no report in the literature as to when and where it was first noticed (Teyssier 1987: 73). Careful bibliographical research we carried out did not allow us to identify either the starting point when palatalization was observed nor its geographical spreading. The first reference to palatalization we found was by Mattoso Câmara 1970: 35) who described the Rio de Janeiro variety. His work was first published in the early seventies and his research was developed in the years preceding publication. He suggests that the palatalization of alveolar stops was then starting to appear in the Rio de Janeiro dialect being that it could be identified as an affrication of the alveolar stop which was more prominent in the primarily stressed position. Abaurre and Pagotto (2002) analyzed data from the seventies to evaluate interdialectal palatalization of alveolar stops in Brazil focusing mainly on geographical distribution. Their findings are illustrated in Table 2.

Table 2: Geographical distribution of palatalization in BP.

Cities	N	%	P.
Recife	66/949	7	.02
Salvador	642/745	85	.73
Rio de Janeiro	841/844	100	.99
São Paulo	543/747	40	.48
Porto Alegre	304/759	59	.18
	2396/4054	59	

Table 2 shows that there are varieties of BP where palatalization is fully implemented (as in Rio de Janeiro) and varieties where palatalization is still incipient (as in Recife). Abaurre and Pagotto (2002: 9) also mention that palatalization is spreading into non-palatalizing varieties and becoming a general feature in BP.

A number of works on palatalization of alveolar stops in BP has been done recently which analyze individually a number of varieties throughout Brazil: Bisol and da Hora (1993) Porto Alegre and Alagoinhas varieties; Mota and Rollemberg (1997) Salvador variety; Souza (1999) and Santiago-Almeida (2000) Mato Grosso variety; Abaurre and Pagotto (2002) Recife, Salvador, Rio de Janeiro, São Paulo and Porto Alegre varieties; Castro and Pisciotta (2002) Paraná, Minas Gerais, Bahia and Sergipe varieties and Cristófaro-Silva (2003) Minas Gerais variety.

In this paper we will consider data from the city of Belo Horizonte. It is a city with a population of 2,23 million, located in the Southeast and it is the 4th major city in Brazil (IBGE <http//www.ibge.gov.com>). Belo Horizonte is a full palatalizing variety of BP (Cristófaro-Silva 2003). We can summarize the facts presented so far by saying that palatalization of alveolar stops triggers an affricate to occur systematically followed by a high front vowel. Belo Horizonte is a fully palatalizing variety where affricates are found in all possible accentual positions and the high front vowel may be oral, nasal or a glide.

In the remainder of this section we will consider two issues related to palatalization of alveolar stops: post-vocalic sibilant palatalization (ˈtɛʃtʃi>ˈtɛʃtʃi *teste* 'test') and lenition of affricates (ˈtɛʃtʃi>ˈtɛʃi *teste* 'test'). These issues are relevant to the understanding of lenition of affricates which will be discussed later.

Let us initially consider palatalization of a post-vocalic sibilant. Post-vocalic sibilants in the Belo Horizonte dialect are systematically alveolar except when the sibilant is followed by an affricate. Thus, post-vocalic sibilants are alveolar word-finally, as in ˈpas *paz* 'peace', and word-internally when followed by any consonant other than an affricate, as in ˈ kaska casca 'peel', ˈaznu *asno* 'donkey', ˈpasta *pasta* 'briefcase'. When a post-vocalic sibilant is followed by an affricate it may be realized as alveolar, as in ˈtɛstʃi and ˈdezdʒi or as alveopalatal, as in ˈtɛʃtʃi *teste* 'test' or ˈdeʒdʒi *desde* 'since'.[2]

Let us now consider the other issue related to palatalization of alveolar stops: lenition of affricates. Lenition involves the weakening and eventual loss of a consonantal segment. The cases we will consider in detail later in this article involve post-vocalic sibilants followed by an affricate where the affricate may be lenited: ˈtɛʃtʃi > ˈtɛʃi *teste* 'test' or ˈdeʒdʒi > ˈdeʒi *desde* 'since' (Barry and Andreeva, 2001).

We would like to point out that by expressing lenition with IPA symbols it appears that a categorical and discrete change takes place, where an affricate is either present or not. We will question this view by showing that a closer examination of acoustic properties of data may offer us important information to understand the transition from one phonetic state to another one in the lenited forms under study.

In the following section we present an experimental analysis to support our claim. The analysis to be discussed in the remainder of this paper will be restricted to unvoiced post-vocalic sibilants followed by affricates: stʃ and ʃtʃ and their corresponding lenited forms. This is due to the fact that voiced post-vocalic sibilants followed by affricates, as in zdʒ and ʒdʒ occur in just a few words. A two hundred thousand word dictionary count showed only nineteen words which present the voiced sequences in question[3]. However, the phenomena that will be discussed below regarding the voiceless sequences of post-vocalic sibilants and affricates are also observed for the voiced sequences.

3. Experimental analysis

3.1. Experimental motivation

Assimilation and lenition are generally described as categorical processes by which a given segmental sequence turns into another one (Lass 1984; Carr 1999). This is witnessed in traditional views of phonology from Generative models to Optimality theory which overtly exclude predictable information of representation. However, a number of recent works has questioned this traditional view by offering insightful analysis of phonological phenomena. These works may focus on perception (Johnson, 1997; Pisoni 1997) or production (Pierrehumbert 2001). It is likely that a full model to analyze phonological phenomena will combine perception and production in an integrated manner (Bybee 2001).

In this section we will focus on the acoustical examination of production. We assume that the reader is familiar with the general principles of acoustic analysis (Kent and Read 1992; Pickett 1999). However, for the sake of clarity we will identify major acoustic correlates of articulatory properties. We first describe the experiment and then we will examine the relevant data.

3.2. The experiment

An experiment was designed to investigate lenition of affricates. Recording took place at the Phonetics Laboratory at the Federal University of Minas Gerais recording studio using a DAT (Digital Audio Tape, Sony TCD D8) and a unidirectional Sony microphone. Participants were told we were testing memory skills through a number of experiments and they agreed to participate in the research project. A number of 8 female and 8 male university students took part. They were all aged 18 to 30 years old. A number of 28 words were analyzed for each one of the participants giving us a total of 443 tokens[4]. Participants were asked to read aloud a set of sentences (cf. appendix). Words to be analyzed were selected according to their stress pattern (stressed or unstressed) and also according to their frequency of token rate: high or low token frequency. Token frequency count was obtained at http://lael.pucsp.br/corpora/index.htm which represents a corpus of

1.182.994 words. The effects of token frequency on the implementation of lenition of affricates will be discussed in section 4. All word tokens were individually edited and analyzed using Praat <www.praat.org> (Boersma and Weenink, 2006).

In the following section we will present the major acoustic characteristics of voiceless stops, affricates and sibilants in order to offer the reader instruments to be able to evaluate the analysis to be presented.

3.3. Acoustic correlates of stops, affricates and sibilants

Let us initially consider the acoustic representation of alveolar stops and affricates. Consider Figure 1 which shows the spectrograms for the words aˈta atar 'to tie up' and aˈtʃa 'made up word for illustrative purposes' focusing on the characteristics of the alveolar stop (on the left) and an alveopalatal affricate (on the right).[5]

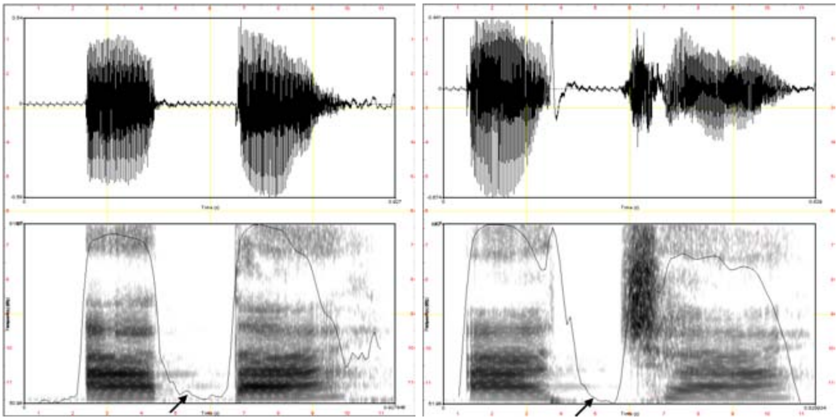


Figure 1: Spectrograms for an intervocalic alveolar stop in the word aˈta atar 'to tie up' and an intervocalic affricate in the word aˈtʃa 'made up word for illustrative purposes'.

The essential articulatory feature of a stop consonant is a momentary blockage of airflow through the vocal tract. Thus, amongst other things, a stop is characterized by very little or no acoustic energy whose spectrographic correlate is the lack of energy. This lack of energy is characterized by a gap or blank area in Figure 1. The gap in the spectrogram reflects the acoustic interval corresponding to the articulatory occlusion during the production of the stop. Another important characteristic of a stop is the drastic drop in intensity (Ladefoged, 1982: 170). We have highlighted the drop in intensity in Figure 1 by an arrow. We would like to call special attention to the drop in intensity as one of the acoustic correlates of a stop since this will be important to the facts to be discussed later.

Notice that the picture on the right handside in Figure 1 represents an affricate. An affricate consists of a stop followed by a sibilant, which in the case of Figure 1 is alveopalatal. Sibilants are acoustically characterized by noise energy which in the picture on right handside in Figure 1 is expressed by the darker area which follows the stop (in the affricate). Notice that there is a rise in intensity when the sibilant occurs after the stop. Thus, we may say that an affricate may be acoustically characterized by two contiguous acoustic properties. For its initial part which corresponds to the stop there is a closure which is represented by the blank space in the spectrogram and also by the drop in intensity. The alveopalatal sibilant in the affricate, which follows the stop, is characterized by the noise energy expressed by the darker area in the spectrogram and the rise in intensity. Let us now consider alveolar and alveopalatal sibilants. Figure 2 presents spectrograms for the words aˈsa *assar* ‘to roast’ and aˈʃa achar ‘to find’.

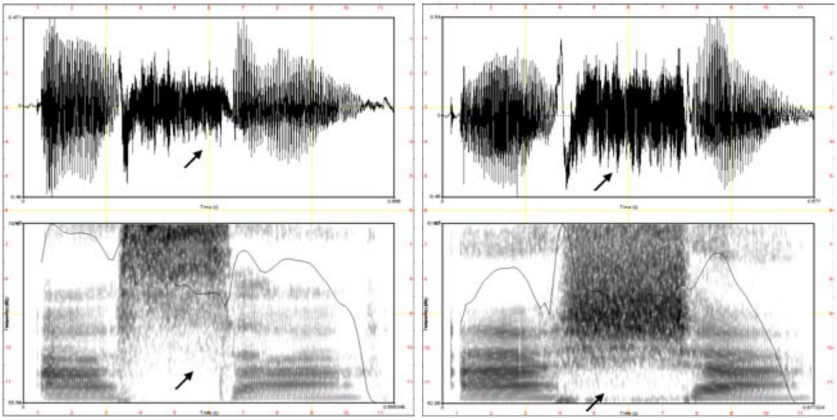


Figure 2: Spectrograms illustrating an intervocalic alveolar sibilant in the word aˈsa *assar* ‘to roast’ and an intervocalic alveopalatal sibilant in the word aˈʃa *achar* ‘to find’.

In both spectrograms in Figure 2 the more prominent darker area characterizes the sibilant which is identified by an arrow. The word aˈsa *assar* ‘to roast’, which presents an intervocalic alveolar sibilant, is illustrated on the left handside in Figure 2. The noise energy from the fricative concentrates over the superior part of the spectrogram. On the right handside it is illustrated the spectrogram for word aˈʃa *achar* ‘to find’ where an alveopalatal sibilant occurs intervocalically. The noise energy from the alveopalatal fricative is also observed and in this case it is observed across the height of the spectrogram. Above the spectrograms are shown oscilograms where we also indicated the sibilants by an arrow. In the following section we will provide an acoustic evaluation of our data in order to consider lenition of affricates.

3.4. Lenition of affricates

One of the crucial issues for theories of sound change is what to consider as a crucial point for a sound change to enter the linguistic system and to progress and settle or not. The debate resolves around two issues. One of them concerns the neogrammarian and lexical diffusion theories (Wang 1969). The other issue concerns the nature of information which is present in lexical representations. We will leave the discussion of neogrammarian and lexical diffusion theories to section 4.

In this section we will address some aspects which are relevant to examine the nature of elements present in lexical representations. We will show that lenition of affricates operates in a gradual fashion being that the acoustic properties inherent to the affricate remain in the signal offering important information to lexical representations (Browman and Goldstein 1992; Albano 2001. See also Kochetov’s contribution to this volume, for a critical analysis of Articulatory Phonology). We suggest that lenition reflects the fact that previous sequential gestures are then manifested simultaneously for at least part of their articulation. Changes in the organization of gestures or their timing produce acoustic-perceptual changes that may be captured in spectrograms. Thus, it is rather the overlapping of articulatory properties that produces the acoustic effect of a gesture being hidden indicating the gradual nature of lenition in BP.

Consider Figure 3 which illustrates post-vocalic sibilants followed by an affricate in the word *constipado* ‘constipated’. An arrow in the spectrogram on the left handside in Figure 3 indicates an alveolar sibilant which is followed by an affricate, i.e., *con[stʃ]ipado*. On the right handside in Figure 3 the arrow indicates an alveopalatal sibilant which is followed by the affricate, i.e., *con[ʃtʃ]ipado*.

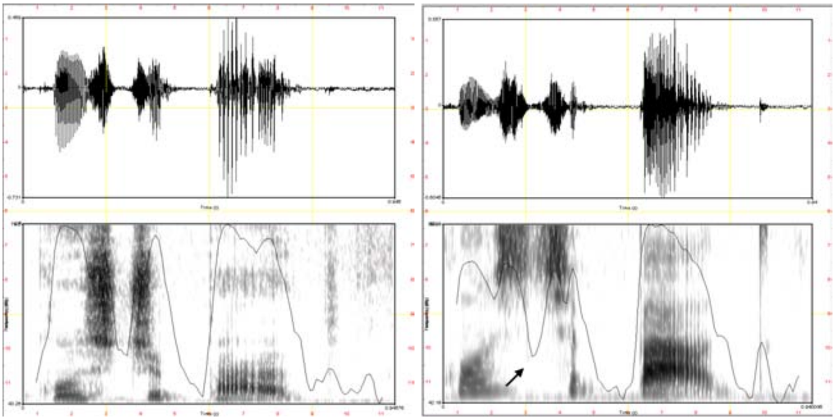


Figure 3: Affricates preceded by a sibilant in the word *constipado* ‘constipated’.

In both spectrograms in Figure 3 one observes the closure of the stop in the affricate which is characterized by a blank area and the drop in intensity. Thus, in both spectrograms in Figure 3 there is a sibilant followed by an affricate. Let us now consider cases which involve lenition. Consider Figure 4.

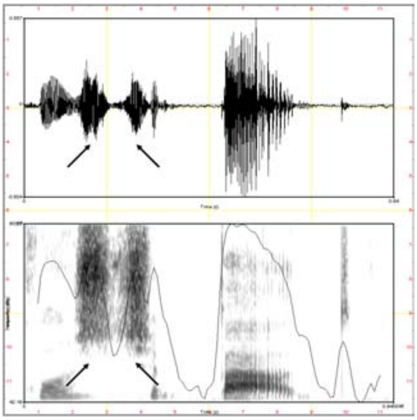


Figure 4: Lenition of an affricate in the word *constipado* ‘constipated’ c*on[ʃʃ]ipado* expressed by a sequence of alveopalatal sibilants with the concomitant drop in intensity.

Figure 4 illustrates the word *constipado* ‘constipated’ where two sibilants occur contiguously and there is no closure between them. The arrows in the spectrogram and in the oscilogram indicate the noise energy corresponding to the sibilants. The drop in intensity between the sibilants may also be observed in Figure 4. However, the drop in intensity which is typical in stops (Figure 1) is not typical in sibilants (cf. Figure 2). We suggest that the drop in intensity which is observed between the sibilants in Figure 4 reflects the former sequential gestures of the affricate closure and the sibilant being manifested simultaneously. In other words it is the overlapping of articulatory properties that produce the acoustic effect of a gesture being hidden. Lenition, thus, in the case under investigation consists of the concomitant articulation which follows from gestural compression. The drop in intensity in Figure 4 expresses a property from the affricate which occurs in the lenited form. Consider now some further cases of lenition which are illustrated in Figure 5.

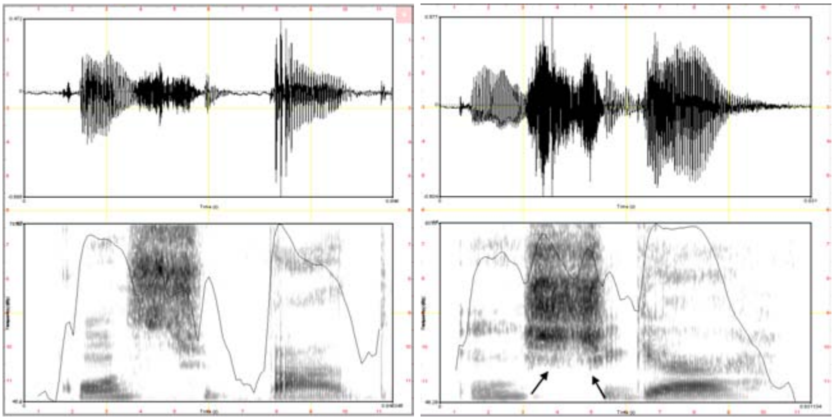


Figure 5: Lenition of affricates in the word *con[sʃ]tipar* ‘constipated’ (on the left) and in the word *in[ʃʃ]tigar* . ‘instigate’ (on the right).

Figure 5 illustrates lenited forms with no closure of the affricate between the sibilants. On the left handside the word con[sʃ]tipar ‘constipated’ shows an alveolar sibilant followed by and alveopalatal sibilant. On the right handside the word in[ʃʃ]tigar) . ‘instigate’ is illustrated showing two alveopalatal sibilants. The arrows in the spectrogram and in the oscilogram highlight the lenited affricates. Notice that in both words illustrated in Figure 5 the drop in intensity between the sibilants which corresponds to the closure in the stop is not as abrupt as was shown on Figure 4. In Figure 5 the drop in intensity is rather smaller. The decrease in intensity follows from greater gestural compression. One may then posit that further gestural compression may yield a regular sibilant to occur. Consider then the spectrogram for the word *con[ʃ]ipado* 'constipated' in Figure 6.

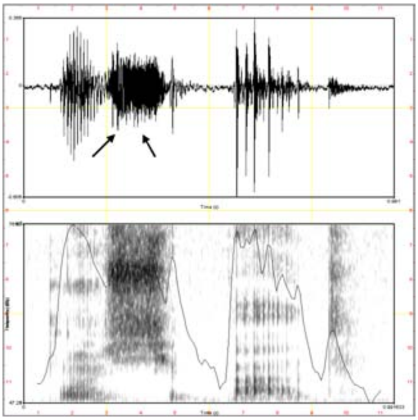


Figure 6: Lenition of affricate in the word *con[ʃ]tipar* '*constipat*ed'.

The spectrogram in Figure 6 illustrates the word 'constipated' where lenition yields the manifestation of an alveopalatal sibilant, i.e. ʃ. Notice that the intensity feature from the stop (in the affricate) is not present during the production of the sibilant shown above. Figure 6 illustrates a case of much greater gestural compression than the ones presented in Figures 4 and 5.

At this point we may say that lenition applies in a phonetically gradual fashion creating a number of patterns that reflect gestural compression which follows from a gestural overlap that takes place gradually. Thus, lenition ranges from the drop in intensity between sibilants which were formerly related to the stop in the affricate to the manifestation of an alveopalatal fricative.

At this stage we can say that there are two sources for alveopalatal sibilants in BP: a regular one (Figure 2) and one related to lenition of affricates (Figure 6). Following a preliminary evaluation of data we posit that the duration of the sibilant is different for each case in question. More specifically we posit the hypothesis that the duration of lenited sibilants is longer than the duration of regular sibilants. Unfortunately, as the experiment we carried out was not specifically designed to examine durational patterns we are not able to pursue this issue here. Sangster (2001) offers and insightful analysis along this line of research which may be taken into consideration in future research.[6] In the following section we consider token frequency effects on the implementation of lenition (Guimarães 2004).

4. Frequency effects on lenition of affricates

In this section we return to one of the core issues related to sound changes, which is the debate between neogrammarian and diffusionist perspectives. In a very concise manner we may say that neogrammarian changes are said to be phonetically gradual and lexically abrupt (Osthoff and Brugmann 1978). Lexical diffusion changes are said to be phonetically abrupt and lexically gradual (Wang 1969). The categorical distinction between neogrammarian and lexical diffusion changes has opened a lively debate (Labov 1994; Oliveira 1991). An additional perspective to neogrammarian and diffusionist theories is suggested by Bybee (2001) who assumes that sound change is both phonetically and lexically gradual. In cases of phonetically motivated sound changes Bybee suggests that more frequently used words are affected first.[7]

In the previous section it was shown that lenition is phonetically implemented in a gradual fashion in BP. In this section we intend to offer evidence that lenition is lexically implemented in a gradual manner through patterns of lexical diffusion.

We examined 28 words which were grouped according to token frequency count: high and low token frequency.[8] We also grouped words in relation to stress, i.e., whether lenition applied to a stressed syllable or to an unstressed one. Results are presented in Table 3.

Table 3: Lenition in relation to stress and token frequency.

	More Frequent				Less Frequent			
	Affricate		Lenition		Affricate		Lenition	
	N.	%	N.	%	N.	%	N.	%
Unstressed	8/176	5%	168/176	95%	41/171	24%	130/171	76%
Stressed	16/48	34%	32/48	67%	33/48	69%	15/48	31%

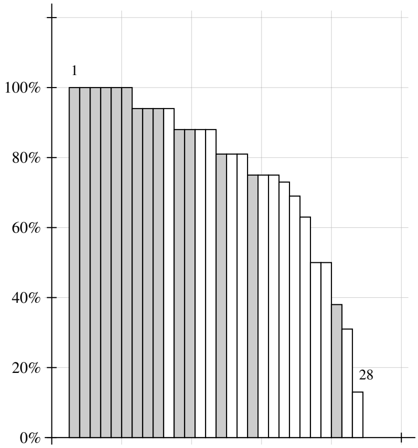
Chi-square = 66.2823178802747. $p \leq 0.001$

We examined whether lenition took place or an affricate occurred. The criterion to identify an affricate was the presence of a blank area in the spectrogram that would characterize the closure for the stop in the affricate. If no gap was observed in the spectrogram the token was then categorized as lenited. We are aware that the discrete categorization of tokens into two distinct groups is not perfectly consistent with the theoretical background we are using. However, we believe that this preliminary analysis may offer insight to future investigations.

We suggest that frequency effects combined with stress patterns may offer us a better understanding of data presented on Table 3. Lenition first appeared in unstressed positions in more frequently used words (95%) then progressed towards unstressed positions in less frequently used words (76%). Thus, lenition is favoured in unstressed positions. In its current stage lenition is progressing towards stressed positions in frequently used words (67%) and lenition is less frequent in stressed positions in rarely used words (31%). These facts show that token frequency and stress are closely related in the implementation of lenition thus supporting Bybee's claim that token frequency is important in the implementation of sound changes. Besides examining token frequency effects we also considered the lexical hypothesis that individual words behave differently towards a given phenomena. A closer examination of lenition rates for each word shows that the general tendency is for more frequently used words to present a higher rate of lenition. Consider Figure 7.

Figure 7 illustrates the rates for lenition in individual words. Notice that more frequently used words have a tendency to present a high rate of lenition, indicated by the grey bars grouped towards the left on Figure 7. On the other hand, less frequently used words have a tendency to present a lower rate of lenition which is indicated by the white bars towards the right in Figure 7. Notice that there are words which do not follow this general expected tendency for token frequency. An example is the word number 10 (*cabalística*) which has a high rate of lenition but is not used that frequently. If we consider the word number 26 (*justiça*) we observe that it has a high token count but a lower rate of lenition. The behaviour of individual words with respect to frequency effects simply indicate that general token frequency tendencies are followed although some lexical items do not behave as expected with respect to token frequency tendencies.

Item		**Token frequency**	**Lenition rate**
1	plástico	1401	100 %
2	característica	758	100 %
3	estatística	672	100 %
4	artística	649	100 %
5	diagnóstico	618	100 %
6	investimentos	4054	100 %
7	instituto	5400	94 %
8	existe	5270	94 %
9	festival	4204	94 %
10	cabalística	10	94 %
11	oeste	4908	88 %
12	destino	1630	88 %
13	nasceste	10	88 %
14	realístico	8	88 %
15	reajuste	3532	81 %
16	ritualístico	9	81 %
17	istingar	10	81 %
18	estilo	3375	75 %
19	pernóstico	6	75 %
20	cipreste	4	75 %
21	pestilento	1	73 %
22	escolheste	2	69 %
23	constipado	6	63 %
24	humanística	8	50 %
25	enrustido	8	50 %
26	justiça	12147	38 %
27	estiva	10	31 %
28	cistite	1	13 %



High frequency words in grey
Low frequency words in white

Figure 7: Frequency effects in lenition in individual words.

The data discussed in this section provides us with evidence for a model which incorporates token frequency effects into the organization of linguistic systems. General tendencies are important but words may have specific behaviour that does not fit into the major token frequency tendencies.

5. Conclusion

This article examined lenition of affricates in BP. It was shown that there is a great range of fine phonetic detail related to the implementation of lenition. The phonetic variability observed is related to the fact that previous sequential gestures are now manifested simultaneously. Changes in the organization of gestures or their timing produce acoustic-perceptual changes that may be captured in spectrograms (Browman and Goldstein 1992; Albano 2001). Thus, an important contribution of this paper is to recognize the emergent character of phonetic categories yielding the transition from one phonetic state to another one. Our results support the view that sound variation and change is implemented in a phonetically gradual fashion.

Additionally, this paper supports the claim that token frequency is relevant in the implementation of sound variation and change. We have shown that more frequently used words undergo lenition at a higher rate than lower frequency ones. We also showed that stress plays an important role in the implementation of lenition of affricates, since words may, if appropriate, behave in a non-predictable manner with respect to frequency effects. Thus, the results presented in this paper support the view that sound variation and change is phonetically and lexically implemented in a gradual manner (Bybee 2000, 2001).

A number of issues raised in this paper still deserves further investigation. The evaluation of durational values of regular and lenited sibilants may contribute to a better understanding of similar segments. Another pertinent issue to be pursued in future works regards the analysis of frequency effects grouped logarithmical manner (Mendoza-Denton et al. 2003). The perceptual evaluation of regular and lenited sibilants also may be investigated. Some preliminary results show that regular sibilants are generally perceived as a sibilant whereas lenited sibilants are mainly perceived as affricates (Guimarães 2004).

Appendix

Sentences used for data collection. Words investigated are underlined

1. O instituto brasileiro de geografia e estatística (IBGE) aponta que, contrariamente ao que se pensa, o número de nascimentos de homens é maior do que o de mulheres em Belo Horizonte.

2. A desconfiança é a característica principal do mineiro. O mineiro da cidade grande tem um estilo de vida agitado. Mas, há muitas vantagens de se morar em Belo Horizonte. Uma delas é o acesso à cultura. O teatro, por exemplo, é uma manifestação artística que vem ganhando

cada vez mais espaço nessa cidade. Além disso, na área musical, um festival de novos talentos tem surgido a cada dia.

3. Belo Horizonte é uma cidade em constante crescimento. Pode-se notar o crescimento desordenado de alguns bairros. O reajuste dos preços do material de construção nos últimos meses parece ter freado um pouco os investimentos no setor da construção civil. O que não é bom nem para quem mexe com construção nem para os possíveis compradores.

4. Uma das atividades que tem crescido é a reciclagem de materiais. Esse trabalho tem mudado o destino de muita gente, que sobrevive de catar papel, plástico e vidro.

5. Com os problemas da Santa Casa, o maior hospital do SUS de Belo Horizonte, muitas pessoas têm ficado sem atendimento. A cistite, uma doença que afeta principalmente as mulheres, tem diagnóstico fácil, mas a falta de medicamentos tem feito avançar esse doença principalmente na população mais carente.

6. Em Venda Nova, uma empregada doméstica acabou ofendendo a secretária do posto de saúde porque teve que esperar horas por um atendimento a seu filho que estava constipado.

7. Um dos grandes problemas que aterroriza a população é a dengue, uma doença transmitida por um mosquito pestilento que se reproduz em locais de água limpa e parada. Em época de chuva, o enchimento de pequenas poças de água contribui para a proliferação desse mosquito.

8. A violência vem crescendo a cada dia em Belo Horizonte. Além do mais, a justiça, como em todos os outros estados, é lenta e falha. A região oeste é uma das mais afetadas pela violência. Recentemente, nesta região, foi fundada uma escola que conjuga ao aprendizado profissionalizante, uma formação humanística.

9. Instituições de combate à droga procuram instigar os jovens a buscar novas atividades que possam preencher o tempo ocioso. Tem-se chegado a conclusão de que não adianta fazer um discurso pernóstico. Os jovens necessitam de ajuda concreta.

10. No dia da fundação dessa instituição, o presidente fez um discurso inicial e, em um tom ritualístico, convocou as empresas particulares a

ajudarem no combate a droga. Um problema que tem mexido com a vida de toda a população.

11. Alguns psicólogos fazem um trabalho voluntário aplicando o chamado teste vocacional para ajudar os jovens que ainda não escolheram a profissão.

12. Os especialistas alertam que mesmo o adolescente quieto e enrustido pode estar precisando de ajuda. Os pais precisam assumir um ponto de vista realístico para encarar os problemas de seus filhos. Existe, em Belo Horizonte, uma instituição de ajuda aos dependentes químicos que oferece cursos aos pais e faz encaminhamentos a clínicas especializadas.

13. A serra do cipó tem uma beleza exuberante. Diversos tipos de plantas são encontrados lá, como cipreste, plantas medicinais e flores. Essa exuberância tem sido prejudicada pelos incêndios, um verdadeiro castigo a natureza.

14. O pai disse a seu filho: "Nasceste e cresceste nesta terra. Agora escolheste outro caminho. Não terás o direito de voltar atrás."

15. Na estiva, foram encontradas muitas caixas de bebidas vindas do México, sem a nota fiscal. Todas foram apreendidas pela polícia.

16. Foi constatado nos últimos meses uma baixa dos preços de roupas no comércio de Belo Horizonte. Um vestido que antes custava R$120,00 pode ser encontrado por R$80,00 ou até menos.

17. O cachimbo era muito usado antigamente. Hoje só usam cachimbo pessoas mais velhas e no interior.

18. Geralmente, nas cidades do interior, ocorre uma disputa cabalística entre dois partidos políticos. Na verdade, levam-se em conta não só as questões políticas, mas também os interesses pessoais.

Notes

1. The authors would like to register support from CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico (grant number: 30.41.21/2002-9) and

CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Research grant). We take the opportunity to thank Jim Scobbie for a careful reading and many insightful suggestions. Of course, the faults are ours. We also take the opportunity to thank the editors and in special Frank Kügler for the careful coordination of the editorial procedure.

2. In Rio de Janeiro all post-vocalic sibilants are alveopalatal: ˈpaʃ *paz* ‘peace’, ˈpaʃta *pasta* ‘briefcase’, haʒga *rasga* ‘rip’ (Callou and Leite 1990). In Recife, a northeastern variety, post-vocalic sibilants are alveopalatal when followed by a coronal consonant: ˈaʒnʊ *asno* ‘donkey’, ˈpaʃta *pasta* ‘briefcase’, ˈdeʒdi *desde* ‘since’, and post-vocalic sibilants are alveolar when followed by noncoronal consoants or followed by a pause: ˈkaska *casca* ‘peel’ or as in ˈpaz *paz* ‘peace’ (Cristófaro-Silva 2001; Mota 2002). We mention here dialectal variation concerning post-vocalic sibilants for the sake of curiosity and as a piece of information for future research on this topic.
3. Michaelis (www.uol.com.br/michaelis). 19 words found with ʒdʒ pattern: *capitisdiminuição, capitisdiminuição, capitisdiminuído, desde, desdita, desditado, desdito, desditoso, desdizer, disdiadococinese, disdiadococinesia, disdiadococinésico, disdipsia, jurisdição, jurisdicionado, jurisdicional, jurisdicionante, jurisdicionar, ormasde.*
4. We expected to collect 448 tokens (16 participants x 28 words). However, some words were read as a different one or the participant did not read it. Then, the total number of analyzed tokens was 443 (missing 5).
5. The figures below illustrate recordings from one of the authors for illustrative purposes. For details of these segments in other languages see: Kent and Read 1992; Johnson 1997; Ladefoged 1982, 2003. The infinitive verbal forms were pronounced without the final r-sound reflecting a characteristic pronunciation in Belo Horizonte see Oliveira 1983.
6. An experiment designed to compare regular and lenited sibilants shall investigate identical or very similar contexts. For example a regular sibilant in the word a[ʃ]e *ache* ‘Find!’ and a lenited sibilant in the word ha[ʃtʃ]e> ha[ʃ]e *haste* ‘pole’. We did not have words with identical environments in our experiment. A preliminary analysis of durational patterns in regular and lenited sibilants is presented in Guimarães (2004) which led us to posit the durational hypothesis presented here. However the number of tokens examined in Guimarães (2004) was very small to offer definite results.
7. Phillips (1984, 2001); Bybee (2001) argue that sound changes which are not phonetically motivated affect less frequently words first.

8. In our experiment low frequency word presented a count below 10 tokens and high frequency words presented a count above 600 tokens (over one million words corpus). An interesting possibility which may be examined in future works is to group token frequency in a logarithmical manner Mendonza-Denton et al. (2003).

References

Abaurre, Maria Bernadete Marques and Emílio Gozze Pagotto
2002 A palatalização das oclusivas dentais no português do Brasil. [Palatalization of dental stops in Brazilian Portuguese]. In *Gramática do Português Falado VII*: novos estudos descritivos; Maria Bernadete Marques Abaurre and Ângela C. S. Rodrigues (eds.) Campinas: Editora da Unicamp.

Albano, Eleonora Cavalcante
2001 *O gesto e suas bordas*: esboço de fonologia acústico-articulatória do português brasileiro. [Gestures and their limit: a preliminary study of Brazilian Portuguese articulatory phonology] Campinas: Mercado de Letras.

Barry, William and Bistra Andreeva
2001 Cross-languages similarities and differences in spontaneous speech patterns. *Journal of the International Phonetic association* 31(1): 51-66.

Bisol, Leda and Dermeval da Hora
1993 Palatalização da oclusiva dental e fonologia lexical. [Palatalization of Dental Stops and Lexical Phonology] *Letras* 5: 25-40.

Boersma, Paul and David Weenink
2006 Praat: doing phonetics by computer (Version 4.4.24) [Computer program]. Retrieved June 19, 2006, from http://www.praat.org/

Browman, Catherine and Louis Goldstein
1992 Articulatory phonology: an overview. *Phonetica* 49: 155-180.

Bybee, Joan
2000 The phonology of the lexicon: evidence from Lexical Diffusion. In *Usage-based models of language*. Suzanne Kemmer and Michael Barlow. (eds.), 65-85. Stanford California: CSLI Publications.

Bybee, Joan
2001 *Phonology and language use*. Cambridge: Cambridge University Press.

Callou, Dinah and Ione Leite
1990 *Iniciação à Fonética e à Fonologia*. [Introduction to Phonetics and Phonology] 8. ed. Rio de Janeiro: Jorge Zahar Editor.

Carr, Philip

1999 *English Phonetics and Phonology: An Introduction.* Oxford: Blackwell Publishing

Castro, Vandersi Sant' Ana and Harumi Pisciotta

2002 Palatização no português do Brasil: dados de atlas linguísticos regionais. [Palatalization in Brazilian Portuguese: data from regional linguistic atlas] *Estudos Lingüísticos 31.*

Cristófaro-Silva, Thaïs

2000 Sobre a quebra de encontros consonantais no português brasileiro. [On the breaking up of consonantal clusters in Brazilian Portuguese]. Estudos Lingüísticos. [Linguistic Studies] 29: 522-527. São Paulo.

Cristófaro-Silva, Thaïs

2001 Fonética e Fonologia do Português: Roteiro de Estudos e Guia de Exercícios. [Phonetics and Phonology of Portuguese] São Paulo: Editora Contexto.

Cristófaro-Silva, Thaïs

2003 Palatalization in Brazilian Portuguese. *Living on the Edge*: 28 papers in honour of Jonathan Kaye. Stefan Ploch (ed.) New York: Mouton de Gruyter.

Docherty, Gerard, Paul Foulkes, James Milroy, Lesley Milroy and David Malshaw

1997 Descriptive adequacy in phonology: a variacionist perspective. *Journal of Linguistics* 33: 275-310.

Guimarães, Daniela Mara Lima Oliveira

2004 Variação sonora nas seqüências de (sibilante + africada alveopalatal) no português de Belo Horizonte. [Sound variantion in (africate + alveopaltal sibilant) sequencies in Belo Horizonte dialect of Brazilian Portuguese]. M. A. dissertation. Federal University of Minas Gerais. Brazil.

IBGE.

2002 Instituto Brasileiro de Geografia e Estatística [Brazilian Institute of Geography and Statistics] <http//www.ibge.gov.com>.

Johnson, Keith

1997 Speech perception without speaker normalisation. In *Talker variability without in speech perception.* Keith Johnson and John W. Mullenix, (eds.), 145-165. San Diego: Academic Press.

Kent, Raymond. D. and Charles. Read

1992 *The Acoustic Analysis of speech.* San Diego: Singular.

Labov, William

1994 *Principles of Linguistic Change: internal factors.* Vol. 1. Oxford: Blackwell.

Ladefoged, Peter

1982 *A course in phonetics.* New York: Harcourt Brace Jovanovich

Ladefoged, Peter
2003 *Phonetic Data Analysis*. Blackwell.

Lass, Roger
1984 *Phonology: an introduction to basic concepts*. Cambridge University Press.

Lingüística Aplicada e Estudos da Linguagem - LAEL
2004 Banco de dados do português. São Paulo: PUC/SP. <www.lael.pucsp.br/corpora>

Mattoso-Camara, Joaquim
1970 *Estrutura da Língua Portuguesa*. [Structure of Portugue Language] Petrópolis: Vozes.

Mendonza-Denton, Norma, Jennifer Hay and Stefanie Jannedy
2003 Probabilistic Sociolinguistics: beyond variable rules. In *Probabilistic Linguistics*, Bod, Rens, Jennifer Hay and Stefanie Jannedy. MIT Press.

Mota, Jacyra A. and Vera Rollemberg
1997 Variantes Africadas Palatais em Salvador. [Variants of Palatal Africattes in Salvador] *Diversidade Lingüística no Brasil*. [Brazilian Portuguese Varieties: <s> in coda in Salvador speakers formal speech] Dermeval da Hora (ed.), 131-140. João Pessoa: Idéia.

Mota, Jacyra A.
2002 *As variantes no português do Brasil. O <s> em coda silábica na norma culta de Salvador*. Ph.D. thesis, Faculdade de Letras, Universidade Federal do Rio de Janeiro.

Oliveira, Marco Antônio de
1983 *Variation and Change in Brazilian Portuguese: The case of the liquids.* PhD thesis. University of Pensylvania.

Oliveira, Marco Antônio de
1991 Controversy Revisited. *International Journal Of The Sociology Of Language*. 89(1): 93-105.

Osthoff , Hermann and Karl Brugman
1978 Preface to morphological investigations in the sphere of the Indo- European languages: 299-209.

Phillips, Betty
1984 Word frequency and the d frequency and the actuation of sound change. *Language* 60: 320-342

Phillips, Betty
2001 Lexical diffusion, lexical frequency, and lexical analysis. In *Frequency and the Emergence of Lingustic Structure*. Joan Bybee and Paul Hopper (eds.), 123-136, Amsterdam: John Benjamins.

Pickett, J. M.
1999 *The acoustics of speech communication*. Boston: Allyn and Bacon.

Pierrehumbert, Janet
2001 Exemplar dynamics: Word frequency, lenition and contrast. In *Frequency effects and the emergence of linguistic structure*. Joan Bybee and Paul Hopper (eds.), 1-19, Amsterdam: John Benjamins.
Pisoni, David
1997 Some Thoughts on "Normalization" in Speech Perception. In *Talker variability without in speech perception*. Keith Johnson and John Mullenix (eds.), 1-8. San Diego: Academic Press.
Sangster, Catherine M.
2001 Lenition of Alveolar stops in Liverpool English. *Journal of Sociolinguistics* 5: 401-412.
Santiago-Almeida, Manuel Mourivaldo.
2000 *Aspectos fonológicos do português falado na Baixada Cuiabana: traços de língua antiga preservados no Brasil (manuscritos da época das bandeiras, século XVIII)*. [Phonological aspects of spoken Portuguese from lowland Cuiabana: preservation of old features from Portuguese language] Ph. D. thesis, Universidade de São Paulo.
Souza, Ulisdete Rodrigues de
1999 *Fonologia do Português Mato-Grossense*: uma perspectiva crioulística.
Souza, Ulisdete Rodrigues de
2000 Portuguese phonology from Mato Grosso: a creole perspective M. A.
Teyssier, Paul
1987 *História da Língua Portuguesa*. [History of Portuguese Language]. Lisboa: Livraria Sá da Costa.
Wang, William. S-Y.
1969 Competing change as a cause of residue. *Language* 45(1): 9-25.
Zsiga, Elizabeth
1995 An acoustic and electropalatographic study of lexical and postlexical palatalization in American English. In *Papers in Laboratory Phonology IV*: Phonology and phonetic evidence. Bruce Connell, Amalia Arvanti (eds.), 282-302. Cambridge: Cambridge University Press.

Silent onsets? An optimality-theoretic approach to French *h aspiré* words[1]

Christoph Gabriel & Trudel Meisenburg

This paper reexamines one of the classical problems of French phonology: the phenomenon of *h aspiré* words (Fouché 1959, Klein 1963, Rothe 1978, Encrevé 1988, Green and Hintze 2004). These vowel initial items display a special behavior with respect to typical phonological processes of French, such as *enchaînement*, *liaison*, and *élision*. Based on the analysis of experimental data we investigate the phonetic realizations of these items and propose an optimality-theoretic (OT) approach in order to account for this phenomenon as well as for the variation encountered in our data.

The paper is organized as follows: We start with an overview of the relevant facts from a rather descriptive point of view (section 1) before presenting the data analyzed for the present study (section 2). While section 3 reviews selected non OT approaches to French *h aspiré* words, section 4 goes into the problem of gradient variation from an optimality-theoretic perspective: We give an overview of the proposals under discussion before putting forward our own account that largely relies on the model of Stochastic OT developed by Boersma and Hayes (2001). Section 5, finally, offers some concluding remarks.

1. The phenomenon

One of the characteristics of French phonology lies within the strong tendency to create regular CV sequences and to avoid hiatus (V.V) through the application of phonological processes such as *enchaînement* (syllabification of a fixed final consonant into the empty onset position of the following word, 1a), *liaison* (surfacing of a latent final consonant into the onset position of the following word, 2a) and *élision* (deletion of certain vowels in pre-vocalic position, 3a). All these processes have in common that they create onsets for vowel-initial words (that would otherwise violate the high-ranked constraint ONSET), but at the same time they mask the left edges of these words by a consonantal element. There is, however, a group of words that – despite of their (semi)vocalic beginning – do not trigger any of these processes and thus rather behave like consonant-initial words (1b and c, 2b and c, 3b and c). Many of them are of Germanic origin (e.g. *HAUNIþA > *honte* 'shame') and until the 16th century they were pronounced with an initial glottal fricative [h] (Thurot 1881: 391-392) which is still reflected in their orthography. This is the reason for the term '*h aspiré* words' used to refer to this group, although other members derive from different sources and do not necessarily display an *h* in their written form (e.g. *le onze* 'the (number) eleven' [lə.ɔ̃z] < lat. UNDECIM).

(1)

a.	*sept amis*	/sɛt/	+	/ami/	→	[sɛ.tami][2]	+*enchaînement*
	'seven friends'	C		V		CV.CV	
b.	*sept garçons*	/sɛt/	+	/gaʁsɔ̃/	→	[sɛt.gaʁsɔ̃]	−*enchaînement*
	'seven boys'	C		C		C.C	
c.	*sept Hongrois*	/sɛt/	+	/ɔ̃gʁwa/	→	[sɛt.ɔ̃gʁwa]	−*enchaînement*
	'seven Hungarians'	C		V		C.V	

(2)

a.	*les amis*	/le(z)/	+	/ami/	→	[le.zami]	+*liaison*
	'the friends'	CV		V		CV.CV	
b.	*les garçons*	/le(z)/	+	/gaʁsɔ̃/	→	[le.gaʁsɔ̃]	−*liaison*
	'the boys'	CV		C		CV.CV	
c.	*les Hongrois*	/le(z)/	+	/ɔ̃gʁwa/	→	[le.ɔ̃gʁwa]	−*liaison*
	'the Hungarians'	CV		V		CV.V	

(3)

a.	*l'abeille*	/la/	+	/abɛj/	→	[la.bɛj]	+*élision*
	'the bee'	CV		V		CV.CV	
b.	*la fille*	/la/	+	/fij/	→	[la.fij]	−*élision*
	'the girl'	CV		C		CV.CV	
c.	*la hausse*	/la/	+	/os/	→	[la.os]	−*élision*
	'the rise'	CV		V		CV.V	

While *h aspiré* words behave like consonant-initial words in the processes illustrated in (1) to (3), the prosodic process shown in (4) – the surfacing of a latent schwa /(ə)/ – applies exclusively to this group.

(4)

a.	*une abeille* ‘a bee’	/yn(ə)/ VC	+	/abɛj/ V	→	[y.na.bɛj] CV	−surfacing /(ə)/
b.	*une souris* ‘a mouse’	/yn(ə)/ VC	+	/suʁi/ C	→	[yn.suʁi] C.C	−surfacing /(ə)/
c.	*une hausse* ‘a rise’	/yn(ə)/ VC	+	/os/ V	→	[y.nə.os] V.V	+surfacing /(ə)/

The facts mentioned so far provide strong evidence for the assumption that the left edge of *h aspiré* words is granted special protection. This protection, however, inevitably yields an increase of syllabic structures that are rather dispreferred in French: While hiatus (V.V as in 2c [le.ɔ̃gʁwa], 3c [la.os], 4c [ynə.os]) only violates ONSET, the constellation C.V (1c [sɛt.ɔ̃gʁwa]) additionally entails a transgression of NOCODA. Such a local conjunction (Smolensky 1993, Tranel and Del Gobbo 2002) or conjoint violation of ONSET&NOCODA is problematic in its phonetic realization as will be demonstrated in the following section. One strategy to avoid these constellations simply consists in not treating *h aspiré* words as such (5b), yielding pronunciations that are rather associated with informal speech styles.

(5)

a. *les haricots* / *quels haricots !* /le(z)/kɛl(z) aʁiko/ → [le.aʁiko]/[kɛl.aʁiko]
‘the beans’ / ‘what beans!’

b. *les haricots* / *quels haricots !* /le(z)/kɛl(z) aʁiko/ → [le.zaʁiko]/[kɛl.zaʁiko]

While some *h aspiré* words seem to lose their special protection, being increasingly treated as regular vowel initial words, other items, such as the numerals and letter words given in (6), insistently retain their special status. There are even new *h aspiré* words emerging from abbreviations (7) and language games such as *verlan* (8).

(6) *le huit* [lə.ɥit] ‘the (number) eight’, *le onze* [lə.ɔ̃z] ‘the (number) eleven’, *la une* [la.yn] ‘the (number) one / front page’, *le R* [lə.ɛʁ] ‘the (letter) R’

(7) *la SNCF* [la.ɛsɛnseɛf] (*Société nationale des chemins de fer*) ‘national railway society’, *le RER* [lə.ɛʁəɛʁ] (*Réseau Express Régional*) ‘local train network’, *les HLM* [le.aʃɛlɛm] (*Habitations à Loyer Modéré*) ‘state-subsidized apartments’

(8) *deux oufs* [dø.uf] (verlan form of *deux fous*) 'tow madmen' (Plénat 1995: 104)

2. Data

As, on the whole, *h aspiré* words display a rather low frequency in the French lexicon, they accordingly do not appear very often in spontaneous speech. In order to get reliable data allowing for a systematic comparison of the speakers' strategies for dealing with these items we conducted an experiment in May 2004 with 12 native speakers of French (exchange students at the University of Osnabrück, Germany).[3] The task, camouflaged as a combined word finding and reading speed test, consisted in responding as fast as possible to a total of 48 visual, mainly written stimuli that appeared on the screen in the same pseudo-randomized order. Subjects directed the pace of the test by hitting the spacebar after reading or verbalizing each stimulus. These included 16 instances of (written) *h aspiré* words, among them *hamburgers* 'hamburgers', *handicapés* 'handicapped', *haricot* 'bean', *harpe* 'harp', *hasard* 'coincidence', *hausse* 'rise', *hautbois* 'oboe', *héros* 'hero', *Hongrois* 'Hungarian', *honte* 'shame' etc. The items appeared in all relevant contexts, i.e. following a fixed consonant, as in (9a), a latent consonant, as in (9b), or a fixed consonant plus latent schwa, as in (9c). We did leave out, however, the constellation final vowel plus *h aspiré* word as in *la honte* or *le hasard*, because the non-elision of [a] and [ə] in the orthographical representation (unlike, e.g., *l'ami* 'the friend') strongly favors the treatment of these items as *h aspiré* words (see also footnote 3).

(9)

a.	*sept Hongrois* 'seven Hungarians'	/sɛt ɔ̃gʁwa/	(fixed C)
b.	*tout Hongrois* 'every Hungarian'	/tu(t) ɔ̃gʁwa/	(latent C)
c.	*une hausse* 'a rise'	/yn(ə) os/	(fixed C + latent schwa)

The recordings were transferred from DAT recorder to computer, transformed into wav sound files, and submitted to an acoustic investigation using Praat speech analysis software. Our attention focused on the length and nature of the interval between the final segment of the preceding item and the initial vowel of the *h aspiré* word: We measured the span between the release of the consonant and the beginning of the vocal cord vibrations or the eventual interruption of the vibrations in the case of a preceding vowel or voiced consonant. Furthermore, we examined the spectral structure of the intervals. The acoustic analyses were constantly checked against our auditory impressions.

The results can be summarized as follows. One group of speakers simply treats a given *h aspiré* word like a usual vowel initial one: These subjects do not leave any interval between the two segments in question (10c) or they produce only a very small one (10a and b, less than 50 ms), thus yielding the impression that the preceding (latent or fixed) consonant is syllabified into the onset position of the *h aspiré* word.[4]

(10)

a. *sept Hongrois* [sɛ.tɔ̃gʁwa] (4/12 speakers)
b. *tout Hongrois* [tu.tɔ̃gʁwa] (4/12 speakers)
c. *une hausse* [y.nos] (1/12 speakers)

The other group does treat the relevant items as *h aspiré* words: Following a fixed coda consonant, the beginning of the vocal cord vibrations is crucially delayed, thus leaving a pause between the preceding coda consonant and the initial vowel, which is occasionally characterized by the phenomenon of creaky voice. Void in some cases, the pause between the two words is often filled with phonetic material caused by constrictions of the glottis. This kind of 'creaky pause' is symbolized in the following as [˷]. In addition, speakers frequently realize a glottal stop, a sound that is usually limited to cases of special emphasis in French. Concerning the stimulus *sept Hongrois*, the duration of the pause is between 80-200 ms.

(11) *sept Hongrois* [sɛt.˷ʔɔ̰̃gʁwa]/[sɛt. ʔɔ̰̃g-]/[sɛt.˷ʔɔ̃g-]/[sɛt.˷ɔ̃g-]… (8/12)

Speakers thus tend to create some sort of 'silent' onset position in order to avoid the C.V constellation that violates both the NoCoda and the Onset constraint. In our data such silent onsets also appear as a strategy to avoid the (less problematic) hiatus, i.e. after a preceding vowel as in *tout Hongrois* (12b; pause 70-160 ms), although the constellation V.V is generally tolerated in French and even regarded as the 'correct' pronunciation in this case (see Fouché 1959: 438-439; 12a).

(12)

a. *tout Hongrois* [tu.ɔ̃gʁwa] (3/12)
b. *tout Hongrois* [tu.˷ʔɔ̃gʁwa]/[tu. ʔɔ̰̃g-]/[tu.˷ɔ̃g-]/[tu.˷ ɔ̰̃g-]… (3/12)

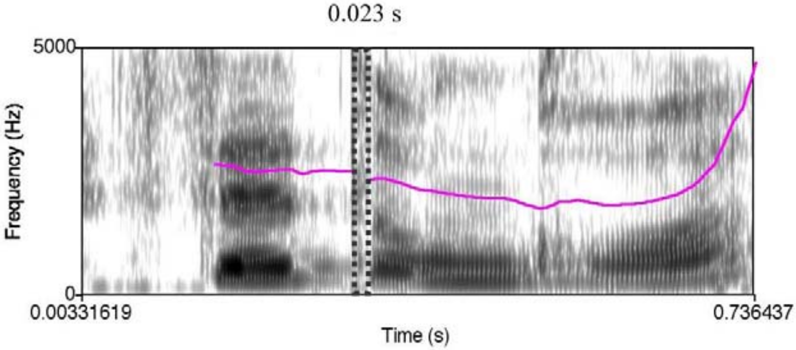
Silent onsets are even created when the underlying form displays a floating schwa that is supposed to surface in order to avoid the problematic C.V constellation yielding the normative hiatus pronunciation *une hausse* [ynə.os] (13a). Concerning this stimulus the inserted pause is between 40-155 ms when schwa surfaces (13b) and between 75-150 ms when no schwa is pronounced (13c).

(13)

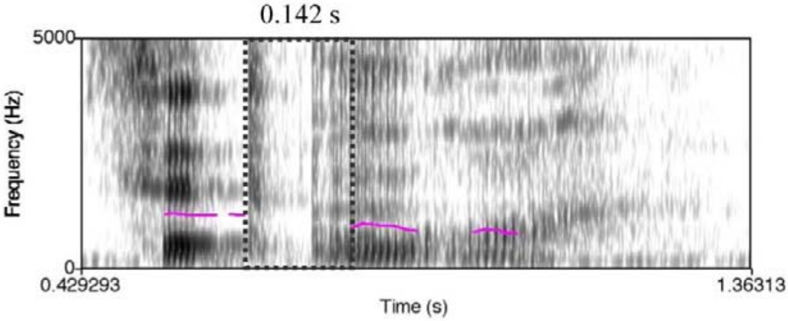
a. *une hausse* [ynə.os] (3/12)
b. *une hausse* [ynə.˷ʔos]/[ynə.˷ʔo̰s]/... (3/12)
c. *une hausse* [yn.˷ʔos]/[yn.˷ʔo̰s]/... (6/12)

The following spectrograms exemplify three different pronunciations of *sept Hongrois*, first treated as regular vowel initial word (14a), then displaying delayed onsets, a rather ‘silent’ one in (14b), and one filled with creaky noise caused by glottal constrictions in (14c). The dotted lined box marks the span between the release burst of the plosive and the beginning of the vowel (equivalent to the interruption of the F0 curve).

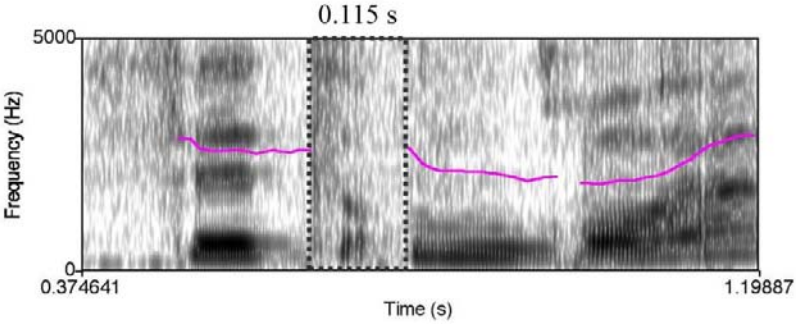
(14)



(a) *sept Hongrois* not treated as an *h aspiré* word [sɛ.tɔ̃gʁwa] (speaker YT).



(b) *sept Hongrois* treated as an *h aspiré* word [sɛt. ʔɔ̰̃gʁwa] (speaker AB).



(c) *sept Hongrois* treated as an *h aspiré* word [sɛt.̰ˀɔ̰̃gʁwa] (speaker BP).

Before elaborating our own OT account we give a short survey of non OT approaches to *h aspiré* words.

3. Non OT approaches to *h aspiré words*

The various attempts that have been made in order to capture the phenomenon of *h aspiré* words are similar insofar as they all rely on a certain kind of lexical marking of the elements concerned. However, the approaches differ in assuming that either a more or less abstract underlying segment at the beginning of the items in question or some inherent structural feature is claimed to be responsible for their special behavior. Scholars belonging to the first group are, among others, Bally (1944: 164) and Klein (1963: 27) who assume a consonantal phoneme zero, Schane (1968: 7-8) who proposes an underlying initial glottal fricative /h/, Dell (1985: 261-262) who departs from an underlying initial glottal stop /ʔ/ and, finally, Pagliano (2003: 638-646), who adopts Dell’s proposal and integrates it into an analysis casted in terms of Charm and Government Phonology (see below for details). One of the representatives of the second line of research is Rothe (1978: 103) for whom *h aspiré* words display a special feature attributing the relevant items to the masculine /lə/ or the feminine /la/ declension rather than to the /l/ declension vowel initial words usually belong to.[5]

Within autosegmental phonology, certain particularities in the skeletal positions of *h aspiré* words and the resulting syllable structures are responsible for the special behavior of the items in question: According to Durand (1986) *h aspiré* words are characterized by word initial empty skeletal slots. Encrevé (1988: 196-203) assumes that they possess an initial skeletal slot that is not filled with melodic material but linked to a syllabic constituent, and that consequently their onset positions are empty but not null. De Jong (1990) presumes that the items in question have no initial skeletal point at all and accordingly lack the onset position of vowel initial words. Charette (1991: 89-90) adopts Dell’s (1985) proposal that *h aspiré* words should be represented as items with an initial constituent dominating a skeletal point but no segmental material and integrates this view into the framework of Government Phonology. Her approach is comparable to Tranel’s (1995) assumption of *h aspiré* words as so-called ‘syllable islands’ that are blocked for certain phonological processes; a necessary prerequisite for these accounts is, once again, diacritic marking of the relevant items in the lexicon.

A recent proposal that relies on both the assumption of an underlying initial consonantal segment /ʔ/ and the concept of special skeletal positions has been put forward by Pagliano (2003: 634-646). Closely following Scheer and Ségéral’s (2001: 117-118) claim that the combined surfacing of schwa and glottal stop at the beginning of *h aspiré* words is restricted to cases of emphasis, she develops a model in which an additional pair of skeletal positions (CV) is inserted into the structure only in case of emphasis, thus allowing for the association of both an (epenthetic) schwa and the underlying glottal stop.[6] From this point of view emphasis is translated by inserting a glottal stop before an initial vowel and, if a glottal stop is already present (i.e. in *h aspiré* words), by inserting a schwa before the (surfacing) segment. Referring to our data Pagliano would claim that there are only two neutral realizations for *une hausse*, either with a surfacing schwa ([ynə.os], as in 13a) or with a glottal stop ([yn.ʔos], as in 13c), whereas the pronunciations given in (13b) that display both features would be restricted to emphatic use.[7] However, the author's judgment concerning the status of a given form as emphatic or not seems to rely more on her native speaker's intuition than on systematic investigation of empirical data. The data we collected do not provide any ground for attributing the combined surfacing of schwa and glottal stop before *h aspiré* words exclusively to emphasis.[8] Until now there seems to be no reliable data supporting the proposal put forward by Pagliano (2003), and we consequently abstain from integrating pragmatic features such as emphasis in the following account.

4. OT account

To our knowledge Tranel and Del Gobbo's (2002) paper is the only recent study that explicitly addresses the problem of *h aspiré* words from an OT perspective.[9] In order to capture the fact that the left edge of *h aspiré* words is granted special protection the authors adopt word specific alignment constraints such as ALIGN-L(*hausse*, σ) militating against the syllabification of (latent or fixed) consonants into the onset position of the relevant item as happens in the case of *liaison* and *enchaînement*. Such an approach, however, crucially challenges the basic OT assumption of the general universality of constraints: A constraint that can exclusively be violated by one and only one lexical item belonging to the lexicon of a certain language, e.g. the French word *hausse*, can hardly be claimed to be universal in the sense that it should be part of the grammar of all languages.

Instead of adopting word-specific constraints that attribute the special behavior of the items concerned to the grammar we propose to pursue a lexicalist approach. In order to account for the phonetic cues we observed at the left edge of *h aspiré* words (glottal stop, pause usually filled with glottal constrictions, creaky vocalic onset, etc.), we assume an underlying initial segment informally symbolized as /ʔ̰/ that embraces all the variation encountered in these surface realizations and is responsible for the special behavior of the items in question. This analysis accounts for the temporal dimension of the delay, which is best captured by the assumption of an underlying consonantal segment.[10] We thus take up the pre OT accounts discussed in section 3 specifying the assumed segment according to the phonetic cues encountered in our data.

As the non-treatment of *h aspiré* words as such cannot systematically be interpreted as a function of a certain speech style, the segment /ʔ̰/ is assumed to be absent from the input of speakers who treat them just like usual vowel initial words, e.g. pronounce *Hongrois* or *hausse* with *liaison* and *enchaînement* as in the examples from our data given in (10). For speakers who switch between these two main realizations, we admit double input forms, e.g. /ʔ̰ɔ̃gʁwa/ and /ɔ̃gʁwa/, from which they choose according to certain pragmatic requirements. The main dimension of variation encountered in the data – i.e. treatment as an *h aspiré* or as a usual vowel initial word – thus constitutes a case of so-called pseudo-optionality (Müller 2003): it is kept out of the grammar and left to the lexicon, in sharp contrast to the proposal put forward by Tranel and del Gobbo (2002).[11]

In order to account for the variation in the output forms resulting from the speakers' different strategies to deal with input forms containing an initial segment /ʔ̰/ we propose an OT account largely based on Boersma's (1998) and Boersma and Hayes' (2001) concept of stochastic OT[12]. Let us first turn to the constraints necessary for dealing with the forms close to the requirements of the norm. Concerning the constellation C.V yielded by the combination of an item displaying a fixed consonant in coda position and an *h aspiré* word, we need a specification of the general ALIGN-L(EFT) constraint stipulating that the left edges of grammatical words coincide with the left edges of prosodic words (15a). Furthermore, a MAX constraint (15b), a markedness constraint banning glottal stops (15c), and the classical DEP-IO constraint are required:

(15)

a. ALIGN(ʔ̰): The left edge of a lexical item beginning with /ʔ̰/ matches the onset of its first syllable.
b. MAX(ʔ̰): /ʔ̰/ in the input is represented in the output.
c. *ʔ: No [ʔ] in the output.
d. DEP-IO: Output segments must have input correspondents.

Given the fact that ALIGN(ʔ̰) occupies a high position in the constraint hierarchy the form [sɛ.tɔ̃gʁwa] is ranked out in tableau (16a). If, on the other hand, the input lacks the initial segment /ʔ̰/, an output containing any of the phonetic cues of a silent onset is ruled out by DEP-IO (16b).

(16)

a. *sept Hongrois* (fixed C + *h aspiré* word)

	/sɛt ʔ̰ɔ̃gʁwa/	ALIGN(ʔ̰)	*ʔ	MAX(ʔ̰)
☞	sɛt._̰ʔɔ̃gʁwa		*	
	sɛ.tɔ̃gʁwa	*!		*

b. *sept Hongrois* (fixed C + vowel initial word)

	/sɛt ɔ̃gʁwa/	DEP-IO	ALIGN(ʔ̰)	*ʔ	MAX(ʔ̰)
	sɛt._̰ ʔɔ̃gʁwa	*!		*	
☞	sɛ.tɔ̃gʁwa				

Let us now turn to the constellation V.V that results from an *h aspiré* word immediately preceded by an item containing a final latent consonant. In order to account for this pattern, we need to split up the constraint militating against glottal stops (15c) into one constraint that bans this segment in intervocalic position (17a) and one that bans it in between consonant and vowel (17b). In addition, we adopt Tranel's (1996) AIF constraint (17c):

(17)

a. *VʔV: No [ʔ] between vowels.
b. *CʔV: No [ʔ] between a consonant and a vowel.
c. AIF: Avoid integrating floaters.

The form [tu.tɔ̃gʁwa] does not respect the special left edge protection typical for *h aspiré* words and is consequently ruled out because of its ALIGN(ʔ̰) violation. The second candidate [tu._̰ʔɔ̃gʁwa] displays a silent onset despite of the fact that the hiatus (V.V) resulting from the non-integration of the floating segment /(t)/ is generally tolerated in French; surfacing of /ʔ/ between two vocalic segments thus constitutes an additional structural effort (violation of *VʔV).

(18) *tout Hongrois* (latent C + *h aspiré* word)

	/tu(t) ʔ̰ɔ̃gʁwa/	ALIGN(ʔ̰)	*VʔV	MAX(ʔ̰)	AIF
☞	tu.ɔ̃gʁwa			*	
	tu._̰ʔɔ̃gʁwa		*!		
	tu.tɔ̃gʁwa	*!		*	*

Whenever an *h aspiré* word is preceded by an item ending in a latent schwa as in *une hausse* the realization of this floating segment corresponds to the norm ([ynə.os]). In order to account for the surface constellation ə.V we assume a constraint requiring the integration of this vocalic element before words with protected left edges, i.e. before underlying /ʔ̰/:[13]

(19) MAX(ə)/_ʔ̰: A floating schwa is integrated before /ʔ̰/.

The relevant tableau is given in (20):

(20) *une hausse* (latent schwa + *h aspiré* word)

	/yn(ə) ʔ̰os/	ALIGN(ʔ̰)	MAX(ə)/_ʔ̰	*VʔV	*CʔV	MAX(ʔ̰)
☞	ynə.os					*
	ynə.˷ʔos			*!		
	yn.˷ʔos		*!		*	
	y.nos	*!	*			*

However, the data presented in section 2 do not only display the kind of variation that can be attributed to different input forms (tableaux 16a vs. 16b), but are also characterized by differing strategies of dealing with forms that contain an underlying initial segment /ʔ̰/. We thus have to account for output forms other than those corresponding to the norm (and emerging as the winning candidates from the evaluation processes depicted in the tableaux 16a, 18, and 20, respectively): Concerning *tout Hongrois* three speakers pronounced [tu.ɔ̃gʁwa] as required by the norm, whereas five participants realized a delayed onset ([tu.˷ʔɔ̃gʁwa]). As for *une hausse* there are even three different ways of dealing with the special protection of the word's left edge: three speakers chose the hiatus solution [ynə.os], two speakers realized an additional silent onset yielding [ynə.˷ʔos], and, finally, six participants left the floating schwa unrealized and assured the left edge protection through silent onset only ([yn.˷ʔos]). As already pointed out in the previous section we do not assume that these slightly differing variants can be attributed in an unambiguous way to different pragmatic contexts. We thus need a model that allows for deriving distinct output forms from one and the same grammar. A relevant proposal concerning this matter comes from Reynolds and Nagy (1994) who define one or more constraints as floating with respect to an otherwise fixed hierarchy. Applying this concept to our data we would need MAX(ʔ̰) as a floating constraint for the case of *tout Hongrois* (21), whereas the various output forms of *une hausse* would require a second floating constraint, Max(ə)/_ʔ̰, that in contrast to the first one moves downwards in the hierarchy (22).

(21)

a. ☞ [tu.ɔ̃gʁwa] Align(ʔ̰) ≫ *VʔV ≫ Max(ʔ̰) ≫ *AIF
b. ☞ [tu.‿ʔɔ̃gʁwa] Align(ʔ̰) ≫ Max(ʔ̰) ≫ *VʔV ≫ *AIF

(22)

a. ☞ [ynə.os] Align(ʔ̰) ≫ Max(ə)/_ʔ̰ ≫ *VʔV ≫ *CʔV ≫ Max(ʔ̰)
b. ☞ [ynə.‿ʔos] Align(ʔ̰) ≫ Max(ə)/_ʔ̰ ≫ Max(ʔ̰) ≫ *VʔV ≫ *CʔV
c. ☞ [yn.‿ʔos] Align(ʔ̰) ≫ Max(ʔ̰) ≫ *VʔV ≫ *CʔV ≫ Max(ə)/_ʔ̰

In order to account for the variation observed we thus have to assume different floating constraints that are promoted and lowered in the hierarchy, respectively. In addition, there are no means to express the relevant frequencies of the forms. A proposal that allows for such frequency predictions comes from Anttila (1997, 2002) who suggests a model of so-called Stratified Grammars (StratG). Such a grammar consists of several strata of constraints that are internally unranked, but strictly ranked with respect to each other. Consequently multiple rankings can be derived from one and the same grammar. Adopting this approach for the analysis of our data we would have to assume the grammar given in (23) displaying a stratum that contains the constraints *VʔV and Max(ʔ̰).

(23) Align(ʔ̰) ≫ { Max(ʔ̰), *VʔV } ≫ …

The following two tableaux can be derived from (23):

(24) *tout Hongrois* (latent C + *h aspiré* word)

stratum

	/tu(t) ʔ̰ɔ̃gʁwa/	Align(ʔ̰)	*VʔV	Max(ʔ̰)	…
☞	tu.ɔ̃gʁwa			*	
	tu.‿ʔɔ̃gʁwa		*!		

	/tu(t) ʔ̰ɔ̃gʁwa/	Align(ʔ̰)	Max(ʔ̰)	*VʔV	…
	tu.ɔ̃gʁwa		*!		
☞	tu.‿ʔɔ̃gʁwa			*	

The frequency prediction resulting from a grammar with a stratum containing two constraints is necessarily 50 % per each output form. In order to account for the different realizations of *une hausse* we have to add two more constraints to the stratum yielding the grammar sketched in (25):

(25) ALIGN(ʔ̰) ≫ { MAX(ʔ̰), *CʔV, *VʔV, MAX(ə)/_ʔ̰ } ≫ …

In (26) we give an example for one of the 24 tableaux that can be derived from the ranking in (25).

(26) *une hausse* (latent schwa + *h aspiré* word)

		stratum				
/yn(ə) ʔ̰os/	ALIGN(ʔ̰)	MAX(ə)/_ʔ̰	*VʔV	*CʔV	MAX(ʔ̰)	…
☞ ynə.os					*	
ynə.˷ʔos			*!			
yn.˷ʔos		*!		*		

(... and 23 tableaux more)

The StratG prediction resulting from the ranking given in (25) is 41.7 % for the forms [ynə.os] and [ynə.˷ʔos], respectively, both being the winning candidate in 10 out of the 24 possible tableaux, and 16.7 % for the form without surfacing schwa ([yn.˷ʔos]) that comes out winning in 4 of the 24 possible tableaux.

Another proposal for an OT grammar that allows for frequency predictions has been made by Boersma (1998) and Boersma and Hayes (2001). In contrast to Anttila's model where such predictions result from the arithmetic proportions determined by the number of constraints contained in a given stratum and the relevant violations, the model developed by Boersma and Hayes is explicitly based on real frequencies and integrates them through stochastic computation into the OT grammar (henceforth StochG). The central assumption of such a StochG model is a continuous ranking scale, on which the constraints occupy fixed ranking values located at more or less distance from one to another. At evaluation time a small amount of noise (according to the standard deviation of 2.0) is added, and the resulting actual ranking values determine the selection points which are relevant for the position of each constraint in a given tableau. As repeated evaluations yield slightly different selection points, a constraint's position on the scale varies around the fixed ranking value. Constraints are thus associated with ranges of values and not with fixed points, and these ranges may overlap to different degrees. The more two neighboring constraints overlap, the higher is the probability that their selection points cross for a given evaluation, the actual tableau thus displaying a reversed ranking. Learnability is achieved by the so-called Gradual Learning Algorithm (GLA), which is fed with learning data and effects the appropriate changes in the ranking values.[14] Gradient variation is thus accounted for by presenting the data under consideration in their relative frequencies.

After feeding the GLA with the frequencies of *tout Hongrois* realized as [tu.‿ʔɔ̃gʁwa] (5/12 speakers) and [tu.ɔ̃gʁwa] (3/12 speakers), a few thousand applications result in the following stochastic ranking values for the relevant constraints (27a) and in the respective distributions of output forms (27b).

(27)

a. Stochastic ranking

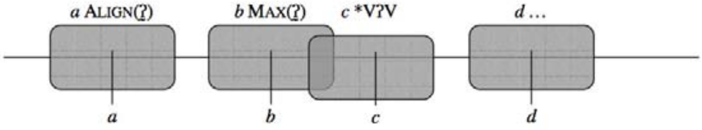
Max(ʔ̰)	100.5
*VʔV	99.5

b. Output distribution

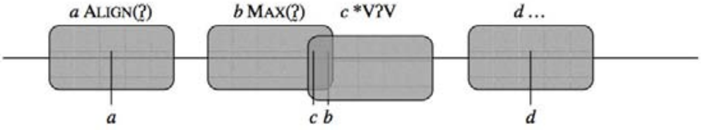
[tu.‿ʔɔ̃gʁwa]	63.2 %
[tu.ɔ̃gʁwa]	36.8 %

The constraint hierarchies that can be derived from (27a) are sketched below: (28a) represents the common result where Max(ʔ̰) dominates *VʔV and [tu.‿ʔɔ̃gʁwa] comes out winning, (28b) represents the less common result with the reverse ranking and [tu.ɔ̃gʁwa] as the winning form.

(28) a. Common result: [tu.‿ʔɔ̃gʁwa] (63.2 %)
Align(ʔ̰) ≫ Max(ʔ̰) ≫ *VʔV ≫ …



b. Less common result: [tu.ɔ̃gʁwa] (36.8 %)
Align(ʔ̰) ≫ *VʔV ≫ Max(ʔ̰) ≫ …



The predicted output distribution for the variation in *une hausse* and the stochastic ranking of the relevant constraints is given in (29):

(29)

a. Stochastic ranking

*VʔV	101.1
Max(ʔ̰)	100.5
*CʔV	98.3
Max(ə)/_ʔ̰	98.3

b. Output distribution

[ynə.os]	27.3 %
[ynə.ʔ̰os]	18.7 %
[yn.ʔ̰os]	54.1 %

If we finally compare the predictions made by the StratG model with those made by the StochG model the latter fit the frequencies displayed in the data much better than the former:

(30)

output forms	our data	StratG prediction	StochG prediction
[tu.ɔ̃gʁwa]	37.5 %	50.0 %	36.8 %
[tu.ʔ̰ɔ̃gʁwa]	62.5 %	50.0 %	63.2 %
[ynə.os]	27.3 %	41.7 %	27.3 %
[ynə.ʔ̰os]	18.2 %	41.7 %	18.7 %
[yn.ʔ̰os]	54.5 %	16.7 %	54.1 %

While the StratG approach does not allow for an adequate modeling of the empirical data, the output distribution calculated by the frequency based StochG approach does. This seems hardly surprising given the fact that in

Boersma and Hayes' model the computation is explicitly based on the frequency distribution found in the relevant data. With the observed frequencies serving as input for the algorithm StochG, however, displays the great advantage to sensibly take into consideration frequency-based differences in the grammar, while such differences are completely ignored in Anttila's approach: StratG relies on purely arithmetic proportions and does not take into account any empirical frequencies.[15] It thus seems to be pure coincidence that Anttila's (1997) computations match the frequencies found in his corpus of Finnish genitive forms quite well.

5. Concluding remarks

French *h aspiré* words and the gradient variation they entail constitute a challenge for phonological theory. As a special type of vowel initial words their main characteristic consists in a protected onset position which is blocked for the syllabification of preceding (fixed or latent) consonants, while preceding latent schwas are supposed to surface. The phonetic cues found in our data exhibit a more or less void but clearly perceptible pause that is eventually followed by a glottal stop and/or a creaky vocalic onset. In order to account for these cues we proposed an underlying initial segment /ʔ̰/, distinguishing *h aspiré* words from regular vowel initial words. But speakers do not always treat *h aspiré* words as such, and these items then behave like usual vowel initial words. To take care of this divergence, we suggested different input forms and a double input for speakers who switch between the two forms. One of the main dimensions of variation found in the data, i.e. the treatment of an item as an *h aspiré* or as a regular vowel initial word, is thus accounted for by different lexical entries starting with /ʔ̰V/ vs. /V/. Variation within the group of *h aspiré* words is determined by the context: Silent onsets are always found when the preceding word ends in a fixed consonant (including the cases in which a schwa does not surface); this constellation always yields C.ʔ̰V. Additional silent onsets optionally occur when the *h aspiré* word is preceded by a regular vowel (yielding V.ʔ̰V) or by a surfacing schwa (yielding ə.ʔ̰V). As has been shown in section 4 this kind of variation that cannot be attributed to different lexical entries or different pragmatic contexts is best accounted for by means of a Stochastic Grammar model that, on the basis of the frequencies attested in the data, replaces the fixed or partly fixed hierarchy by the concept of overlapping constraints.

Notes

1. Earlier versions of this paper have been presented and discussed at several occasions, among others at *Phonetik und Phonologie 1*, Potsdam 2004, at *OCP2* and at the PFC conference *Phonological variation: the case of French,* both in Tromsø, Norway, 2005. We would like to thank all our colleagues for their helpful feedback and comments. Special thanks go to Bill Barry (Saarbrücken) for his support in the analysis of the sound files and to Brechtje Post (Cambridge), Maria Selig (Regensburg), and Caroline Féry (Potsdam) who thoroughly reviewed the prefinal version of our text. The usual disclaimers apply.
2. The notation of underlying forms is simplified and does not entail any theoretical assumptions concerning issues other than the problem addressed here. Syllable boundaries are indicated only at the site under discussion.
3. We also scrutinized the recordings that we made within the context of the PFC project (*Phonologie du Français Contemporain*; see Durand et al. 2002 and http://www.projet-pfc.net) in February/March 2002 in Lacaune (Tarn, Southern France), but there were only few relevant data: *le hasard* ‘the coincidence’ from the fictitious newspaper article that all our 13 speakers realized – as expected – with a surfacing schwa yielding the hiatus pronunciation required by the norm [lə.azaʁ] (V.V), plus three occurrences of *h aspiré* words in the transcribed parts of the interviews: *une haie de buis* [ynə.ɛdəbɥi] ‘a box-hedge’, realized with a surfacing schwa as required by the norm (as in 4c) and two occurrences of *Hongrie* ‘Hungary’ and *hongroise* ‘Hungarian (f.)’, respectively, pronounced by one and the same speaker, but displaying variation insofar as *en Hongrie* ‘in Hungary’ is realized without *enchaînement*, yielding the output form [ɑ̃.ɔ̃gʁi] (as in 2c), whereas *hongroise* in *d’origine hongroise* ‘of Hungarian origin’ is treated as a regular vowel initial word, its left edge consequently being masked through application of *enchaînement* [dɔʁiʒi.nɔ̃gʁwazə] (as in 5b). Given the sparse results these data will not be considered any further in this paper.
4. Note that this holds as well for *quel héros !* ‘what a hero!’ which is pronounced [kɛ.leʁo] by 6 speakers (= 50%), in contrast to Tranel (1995: 812) who explicitly rules out [kɛ.leʁo] in favor of [kɛl.eʁo].
5. From this point of view the situation in French would be comparable to Latin where the type of declension class cannot necessarily be inferred from the morphological shape of the noun.
6. Without this insertion of additional skeletal positions through emphasis only one of the two elements, i.e. either schwa or the glottal stop can surface because of the particular government relations assumed within the Charm and Government model. For further details concerning this framework see the basic papers by Kaye, Lowenstamm and Vergnaud (1985, 1990) and the relevant chapters in Pagliano (2003).

7. It should be pointed out that for Pagliano (2003: 635-636) all pre-*h aspiré* schwas are epenthetic, even those corresponding to a feminine morpheme as in *grosse housse* [gʁosə.us] ‘thick bedspread’. Furthermore she assumes – contra Dell (1985: 186) – that schwa epenthesis can apply in contexts such as *quel hêtre* ‘what a beech’ yielding [kɛ.ləɛtʁ(ə)], an adequate output form claimed to be on a par with [kɛl.ʔɛtʁ(ə)]. If so, there should be occurrences of epenthetic schwa in utterances like *sept Hongrois* or *quel héros*, yielding output forms such as [sɛtə.ɔ̃gʁwa] or [kɛlə.eʁo] that are, however, unattested in our data. The same holds for Green and Hintze’s (2004) corpus.
8. Even if we assumed that some of our speakers used their ‘emphatic speech register’ in the experiment, there would be hardly an explanation for the fact that the same speakers used the emphatic pronunciation with one stimulus and a neutral one with others.
9. Tranel (1996: 442-446) gives a first sketch of how *h aspiré* words could be treated in an OT framework, positing with Align-left ≫ Onset a special ranking – and consequently an extra grammar – for this group of words (vs. Onset ≫ Align-left otherwise). Tranel (2000) and Féry (2003) thoroughly address the different aspects of *liaison* and analyze the resulting syllable structures but do not consider *h aspiré* words.
10. To proceed on the more abstract assumption of an underlying void is further away from surface realizations and unnecessarily complicates the analysis.
11. Another possible solution can be seen in assuming that initial /ʔ̱̰/ becomes a floating segment whenever a given *h aspiré* word loses its specific left edge protection. This would suggest extending the idea of a floating segment to all vowel initial words. Such a generalized beginning with /(ʔ̱̰)/ would also account for the fact that regular vowel initial words can be realized with a glottal stop in emphatic contexts, e.g. *c’est incroyable !* [setʔɛ̃krwajabl] (Pagliano 2003: 621). It should be pointed out, however, that this exclusively holds for lexical categories, given the fact that function words such as vowel-initial pronouns or auxiliaries hardly undergo glottalization, even in emphatic contexts. For reasons of space this approach will not be discussed any further here. The same holds for the assumption of an underlying underspecified consonant [+cons, -voc], as proposed by Caroline Féry (p.c.).
12. For an application of this model to German data see Féry et al. (this volume).
13. See Boersma (2007) for a different proposal.
14. The required computations can be made with the software Praat in which the GLA is integrated.
15. It should be emphasized that the frequency predictions made by StochG entirely result from statistic projections based on the data the algorithm is fed with. The relevant computations thus represent a modeling of the data distribution rather than a prediction concerning the dynamics of language change.

References

Anttila, Arto
1997 Deriving variation from grammar. In *Variation, Change, and Phonological Theory,* Frans Hinskens, Roeland van Houten, and Leo Wetzels (eds.), 35-68. Amsterdam: Benjamins.

Anttila, Arto
2002 Variation and Phonological Theory. In *The Handbook of Language Variation and Change,* Jack K. Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds.), 206-243. Malden, Mass.: Blackwell.

Bally, Charles
1944 *Linguistique générale et linguistique française.* Bern: Francke.

Boersma, Paul
1998 *Functional Phonology.* Ph.D. thesis, Amsterdam: Universiteit van Amsterdam.

Boersma, Paul
2007 Some listener-oriented accounts of h-aspiré in French. *Lingua* 117: 1989-2054.

Boersma, Paul and Bruce Hayes
2001 Empirical Tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32: 45-86.

Charette, Monik
1991 *Conditions on phonological government.* Cambridge: Cambridge University Press.

De Jong, Daan
1990 On floating consonants in French. In *Proceedings of the 21st Western Conference on Linguistics.* El Paso: University of Texas.

Dell, François
1985 *Les règles et les sons.* Paris: Hermann (2nd ed. 1973).

Durand, Jacques
1986 French liaison, floating segments and other matters in a dependency framework. In *Dependency and Non-Linear Phonology*, Jacques Durand (ed.), 161-201. London: Croom Helm.

Durand, Jacques, Bernard Laks, and Chantal Lyche
2002 La *Phonologie du français contemporain* (PFC): usages, variétés et structure In *Romanistische Korpuslinguistik: Korpora und gesprochene Sprache,* Claus Pusch and Wolfgang Raible (eds.), 93-106. Tübingen: Narr.

Encrevé, Pierre
1988 *La liaison avec et sans enchaînement. Phonologie tridimensionelle et usages du français.* Paris: Seuil.

Féry, Caroline
2003 Liaison and syllable structure in French. Ms., University of Potsdam.

Féry, Caroline
this volume Gradient dorsal nasal in Northern German.

Fouché, Pierre
1959 *Traité de prononciation française.* (2nd ed.) Paris: Klincksieck.

Green, John N. and Marie-Anne Hintze
2004 Le *h aspiré* en français contemporain: stabilité, variation ou déclin? In *Variation et francophonie*, Aidan Coveney, Marie-Anne Hintze, and Carol Sanders (eds.), 241-280. Paris: L'Harmattan.

Kaye, Jonathan D., Jean Lowenstamm, and Jean-Roger Vergnaud
1985 The internal structure of phonological representations: A theory of Charm and Government. *Phonology Yearbook* 2: 305-328.

Kaye, Jonathan D., Jean Lowenstamm, and Jean-Roger Vergnaud
1990 Constituent structure and government in phonology. *Phonology* 7: 193-231.

Klein, Hans-Wilhelm
1963 *Phonetik und Phonologie des Französischen.* München: Hueber.

Müller, Gereon
2003 Optionality in optimality-theoretic syntax. In *The Second* Glot International *State-of-the-Article Book. The Latest in Linguistics*, Lisa Cheng and Rint Sybesma (eds.), 289-321. Berlin: Mouton de Gruyter.

Pagliano, Claudine
2003 *L'épenthèse consonantique en français. Ce que la syntaxe, la sémantique et la morphologie peuvent faire à la phonologie : parles-en de ta numérotation ?impossible.* Ph.D. thesis, Nice: Université de Nice.

Plénat, Marc
1995 Une approche prosodique de la morphologie du verlan. *Lingua* 95: 97-129.

Reynolds, William and Naomi Nagy
1994 Phonological Variation in Faetar: An Optimality Account. In *Papers from the Parasession on Variation and Linguistic Theory* (CLS 30), Katherine Beals et al. (eds.), 277-292. Chicago: The Linguistic Society.

Rothe, Wolfgang
1978 *Phonologie des Französischen.* Berlin: Schmidt.

Schane, Sanford A.
1968 *French phonology and morphology.* Cambridge, Mass.: MIT Press.

Scheer, Tobias and Philippe Ségéral
2001 La coda-miroir. *Bulletin de la Société de Linguistique de Paris* 96: 107-152.

Smolensky, Paul
1993 Harmony, Markedness, and Phonological Activity. Handout to talk presented at Rutgers Optimality Workshop 1, October 23, New Brunswick, N.J. [ROA-87].

Thurot, Charles
1881 *De la prononciation française depuis le commencement du XVIe siècle, d'après les témoignages des grammairiens.* Tome 1. Paris: Imprimerie nationale. Réimpression Genève: Slatkine [1966].

Tranel, Bernard
1995 Current Issues in French Phonology: Liaison and Position Theories. *The Handbook of Phonological Theory*, John A. Goldsmith (ed.), 798-816. Cambridge, Mass.: Blackwell.

Tranel, Bernard
1996 French Liaison and Elision Revisited: A unified account within Optimality Theory. In *Aspects of Romance Linguistics*, Claudia Parodi et al. (eds.), 433-455. Washington, D.C.: Georgetown University Press.

Tranel, Bernard
2000 Aspects de la phonologie du français et la théorie de l'optimalité. *Langue française* 126: 40-71.

Tranel, Bernard and Francesca Del Gobbo
2002 Local Conjunction in Italian and French Phonology. In *Romance Phonology and Variation*, Caroline R. Wiltshire and Joaquim Camps (eds.), 191-218. Amsterdam: Benjamins.

Gradient dorsal nasal in Northern German[1]

Caroline Féry, Constance Hohmann & Katharina Stähle

1. Introduction

There has been in the past a large agreement among researchers of spoken language that phonetics studies gradient facts whereas phonology investigates categorical phenomena. It is true that phonologists tend to elaborate tidy inventories of phonemes, syllables or accent patterns, whereas the distribution of the really existing consonants in the consonantal space or of the vowels in the vocalic space are more variable (see Cohn 2006, Keating 1987, Ohala and Ohala 1995 among others). But the distinction between categorical phonology and variable phonetics has been questioned in the last decades, and an increasing number of researchers investigate the claim that phonology should be able to account for variation and gradience (see for instance Pierrehumbert 2003, Pierrehumbert, Beckman, and Ladd 2000, Frisch 2000, Boersma and Hayes 2001, Bybee 2000a,b, 2001, 2003, Cohn 2006 and many others).

If this view is adopted, gradience can be said to originate at different places in phonology. One of the main reasons for gradience is that language continuously evolves and successive generations or groups of individuals may have slightly different grammatical systems from their parents' or neighbors' ones (see for instance Jacewicz et al., this volume). It may also be the case that some social groups implement the changes earlier or more readily than others (see among others Labov 1994, Bybee 2003, Kiparsky 2000, Kroch 2001 and Šimáčková, this volume). As a second source of gradience, categories are not always invariant: the acoustic and articulatory correlates of 'continuant' and 'stop' obstruents, for instance, can be gradient according to context or dialect. Spirantization can also emerge as a lenition effect which is more or less completed, depending on the environment (see for instance Bybee 2001 and Kirchner 2001 for lenition). Our articulatory organs are such that some segments may be unambiguously realized in some contexts, but be more variable or resemble other segments in different contexts (see Guimarãrens and Cristófaro-Silva, this volume, for an example). Neutralization of contrasts between voiced and unvoiced segments is achieved in some environments but not in others, as the acoustic cues differ (Steriade 2001, Darcy et al., this volume). A third source of variation is to be found in the application of allophonic rules. In some cases, distinct categories are targeted, as in t/d deletion in English, or the dorsal nasal in German investigated in this paper. T/d is deleted in English as a dependent variable of morphological category, syllable structure, prosodic phrasing and so on. To mention some examples, it is more frequently deleted in words like *West* or *breakfast*, where the final coda is complex and does not stand for an inflectional morpheme than in *played* and *fed*, where the coda is simple and is morphemically meaningful. It will be shown in this paper that the emergence of [k] after a dorsal nasal in German, the subject of investigation in this paper, is also influenced by a number of phonological factors, like the prosodic category that the word containing [ŋ] ends, the presence of an accent, and the quality of the vowel preceding the dorsal nasal. In short, phonological gradience is not random. Rather the factors weighing on the variation are to be made part of the grammatical model responsible for the variation.

If it is true that phonology is gradient as we think it is, we need a grammar which can account for the gradient effects. It seems that standard generative theories, despite many deep insights about grammatical mechanisms, are ill-equipped to account for such phenomena because the output of re-write rules or of standard constraint interactions is categorical. There have been a number of proposals in classical generative phonology of how the outputs of grammar can be variable, like co-phonologies as proposed by Anttila 2001, or stochastic approaches in OT (Boersma 1998, Boersma and Hayes 2001) as well as older accounts along the line of variable transformational rules (Sankoff and Labov 1979). It will be shown in this chapter that a stochastic optimality theoretic approach is appropriate for the data investigated. But first, the phenomenon is introduced in a derivational framework, as well as the experimental results.

2. Alternation between [ŋ] and [ŋk] in a derivational perspective

In the generative tradition of German phonology, the emergence of [ŋ] and [ŋk] has been accounted for with the help of ordered derivational rules which change an underlying form into a surface one. Isačenko (1963) was the first to propose that some instances of the dorsal nasal are the consequence of a derivation from the sequence /n+g/, the remaining ones being the result of an assimilation to a following voiceless dorsal stop, as in *Bank* 'bank' or *schenken* ‘to give as a present’. This insight has been adopted and further developed by many German phonologists (see for instance Wurzel 1980, Hall 1992, Ramers and Vater 1992, Yu 1992, Wiese 1996, Ito and Mester 2003 and Féry 2003). In a linear derivational model, Nasal Assimilation, Final Deletion and g-deletion are ordered rules, which apply successively in a given order, one rule providing a new structural description, to which the next ordered rule may or may not apply. In the standard dialect of German, Nasal Assimilation applies first and *g*-deletion second. Nasal assimilation changes a coronal nasal into a dorsal one before a dorsal stop, and *g*-deletion deletes /g/ after a dorsal nasal in some environments. A third rule, Final Devoicing, cannot apply because it is ordered after *g*-Deletion, at a stage of the derivation at which its structural description has been bled by application of *g*-deletion. At this point, the [g] to be devoiced is no longer present.[2]

The derivation of the words *Bank* ‘bank’ and *lang* ‘long’ in the standard dialect is illustrated in (1).

(1) Derivation of *Bank* and *lang* in Standard German:

/bank/	/lang/	
baŋk	laŋg	Nasal Assimilation
—	laŋ	*g*-deletion
—	—	Final Devoicing
[baŋk]	[laŋ]	

In the Northern variety, by contrast, the order of application of Final Devoicing and *g*-deletion is reversed. Final Devoicing applies just after Nasal Assimilation, and it is now *g*-deletion which cannot take place because its structural description has been bled. This derivation is illustrated in (2). As a result, *Zeitung* is pronounced [t^{s}ai̱̯tʊŋk]. Notice that *Bank* is pronounced the same in both dialects, since there, only Nasal Assimilation takes place, and *k* is not deleted.

(2) Derivation of *Bank* and *Zeitung* in Northern German:

/bank/	/t^{s}aɪ̯tʊng/	
baŋk	t^{s}aɪ̯tʊŋg	Nasal Assimilation
—	t^{s}aɪ̯tʊŋk	Final Devoicing
—	—	*g*-deletion
[baŋk]	[t^{s}aɪ̯tʊŋk]	

Though interdialectal variation is well accounted for in the derivational approach, variation *in a single dialect* is not. The rules apply categorically, as soon as their structural description is met, and their order of application is not subject to change. According to the rule sequences in (1) and (2), all occurrences of /ng/ are realized phonetically as [ŋ] in the standard dialect, and all occurrences of this sequence emerge as [ŋk] in the Northern dialect. But, as will be documented in the next section, this categoriality does not correspond to reality. Instead, we also find variation in the dialects themselves.

Subsequent generative approaches, like the autosegmental framework, or the standard OT model, do not solve the problem caused by the strict categoriality of the derivational rules. Both the representational model and the optimality-theoretic framework predict a similar invariant application of phonological processes. In a nonlinear pattern (see for instance Hall 1992 and Wiese 1996 for autosegmental accounts of nasal assimilation and *g*-deletion), the place of articulation of the stop spreads regressively, which causes the underlying coronal nasal to become dorsal. After having triggered assimilation, the whole segment slot is deleted, which is expressed by delinking of the major features of this segment or the skeletal slot. Though the format of the rule has changed drastically, the spirit has not. As soon as the structural description of a rule is met, the rule applies and a categorical structural change is the result.

In standard OT, constraints on the surface representation of phonological forms determine which one among a large set of candidates fulfills the constraints best (or violates them least). Constraints are universal, violable and ranked, and the evaluation of the candidates applies in parallel. We return to an optimality-theoretic model of the nasal dorsal in section 4.

3. Experimental study

When listening to a speaker of Northern German, it is conspicuous that the categoriality assumed in standard generative accounts is not supported by the data. Speakers realize both [ŋ] and [ŋk] in a way that can first appear as random. However, half a century of quantitative phonological research, starting with Labov (1966), has taught us that the impression of randomness may be an illusion, and that a precise survey of the data according to relevant parameters may considerably reduce the freedom of variation. It could even be the case that, if it is possible to pin down all factors playing a role in the choice of one allomorph over the other, we find ourselves in a deterministic situation in which variation is completely eliminated. With this idealized hypothesis in mind, an experiment was conducted in order to identify the factors bearing on the choice of one or the other variant in Northern German.

3.1. Experimental set-up

In designing an experiment for studying the factors influencing the distribution of [ŋ] and [ŋk] in Northern German, efficiency and economy were aimed at, in the manner of sociolinguistic surveys. The experimental set-up had to meet several criteria. It was intended to elicit spontaneous but comparable material, a consideration leading to the elimination of spontaneous speech. It was also crucial that the speakers would utter a reasonable number of words containing a dorsal nasal within a short period of time. Instead of using completely free and thus random speech, an experiment was elaborated in which the speakers narrated a story in their own way, but in which some of the lexical material was imposed. To meet these aims, we wrote a short story which contained a large number of words with [ŋ] or [ŋk]. The informants first read the story silently, and in a second step told the story in their own way and rhythm. Recordings were made in a quiet room on a DAT tape-recorder, after the experimenter had given initial instructions as to the procedure of the session. The speakers were naive as to the aims of the recordings. To make sure that they would use the intended words, accompanying drawings were presented. There were also some rare interventions and questions from the experimenters in order to elicit a missing word. The story telling took between 2 minutes 26 seconds and 3 minutes 49 seconds, depending on the speaker. Altogether 18 minutes 48 seconds of speech were analyzed, containing 201 occurrences of words ending with a dorsal nasal.

3.2. Subjects

The results of six informants were evaluated, three men and three women from Potsdam or from the East of Berlin who had lived there all their life. Except for gender, the speakers were quite homogeneous in their social class and age. They were from the middle class and were between 31 and 37 years old. They were friends or relatives of the experimenters. All informants spoke a relatively standard language with only few other dialectal characteristics.

3.3. Analysis

The recordings were first evaluated auditorily by the experimenters. They decided on the basis of what they heard whereas the sound produced was [ŋk], [ŋ] or an intermediate realization with a lenited [k]. In a second step, the first author as well as a fourth, independent listener reevaluated the data, partly auditorily and partly on the basis of spectrograms. These two latter judges classified all items in only two categories, [ŋk] or [ŋ]. They agreed on every single realization, and the evaluation presented below is based on this second evaluation. It must be mentioned at this place that the realization involved a great deal of variation, with many lenited stops. However, a categorical decision as to whether the stop was there or not, was always available, and this guiding line was used for the final decision. At this place an important difference between t/d deletion in English and [ŋk] in German should be mentioned. According to Bybee (2001), t/d deletion is an historical process, whereas [k] is not in the process of being deleted in the phonology of German. Rather, if anything, it is an epenthetic segment, or an excrescent one.

The complete list of words uttered by the informants is reproduced in (3), given in tokens, together with the number of their allophones.[3]

(3)

	Word	Gloss	Total	[ŋ]	[ŋk]
[a]	Wolfgang	‘name’	25	13	12
	Hang	‘slope’	3	3	0
	Gang	‘walk’	3	1	2
	schwang	‘swang’	1	1	0
	klang	‘sounded’	1	1	0
	lang	‘long’	12	3	9
	gelangweilt	‘bored’	3	2	1
	Fischfang	‘fish catch’	6	6	0
	langweilig	‘boring’	5	4	1
	entlang	‘along’	4	1	3
	sprang	‘sprang’	5	5	0
	Gesang	‘singing’	3	1	2
	sang	‘sang’	2	2	0
[ʊ]	Wohnung	‘appartment’	8	5	3
	Ahnung	‘presentiment’	2	1	1
	Tageszeitung	‘newspaper’	2	1	1

continue

	Word	Gloss	Total	[ŋ]	[ŋk]
	Mietwohnung	‘appartment for rent’	2	0	2
	Kreuzung	‘crossing’	5	2	3
	Achtung	‘attention’	5	3	2
	Zeitung	‘newspaper’	4	0	4
	Umleitungsschild	‘detour sign’	1	1	0
	Begrüßung	‘greeting’	6	2	4
	Umleitung	‘detour’	6	0	6
	Beschilderung	‘sign’	2	0	2
	Wohnungstr	‘door of the flat’	1	1	0
	Wohnungsannoncen	‘newpaper advertisement’	1	1	0
	Wohnungsinserate	‘newpaper advertisement’	1	1	0
	Wohnunganzeigen	‘newpaper advertisements’	1	0	1
	Zeitungsinserate	‘newpaper advertisement’	1	1	0
	Richtung	‘direction’	3	3	0
	Unterhaltung	‘entertaining’	1	1	0
[ɛ]	Peng	‘name’	52	52	0
	eng	‘narrow’	3	3	0
	ging	‘went’	7	7	0
[ɪ]	Belling	‘name’	13	12	1

In a second step, a phonological analysis was performed. The recordings were transcribed and Phonological Phrase and Intonation Phrase boundaries were indicated. Breaks in the timing were also localized. Word-medial and word-final occurrences of [ŋ] and [ŋk] were divided into two classes, as mentioned before. Decisions were also taken as to the accent status of the words.

3.4. Results

The main factors bearing on the distribution of [ŋ] and [ŋk] belong to a small group of phonological properties: the boundaries of prosodic domains, the quality of the preceding vowel and the accent, both individually and in relation with each other. Although number of syllables, morphemic structure (root, stem, word) and lexical categories (nouns, adjectives, verbs and preposition) were also examined, no effect was found for these factors. Taking all word final dorsal nasals together, the distribution indicates a clear preference for a realization without a voiceless stop (141 of 201 or 70%), as indicated in (4). In 60 cases (or 30% of the total realizations), a voiceless stop was present. Table (4) shows first the distributions of the two variants for each speaker separately. It can be noticed from this table that the speakers show a wide range of variation as to the proportion of dorsal nasals with or without voiceless stop (between 17% and 42% of their total occurrences), though no speaker had more [ŋk] than [ŋ]. No gender specific tendency is noticeable.[4]

(4) Contextless distribution

	[ŋ]	[ŋk]	Total
Speaker 1 (f)	34 (79%)	9 (21%)	43
Speaker 2 (m)	26 (65%)	14 (35%)	40
Speaker 3 (f)	19 (58%)	14 (42%)	33
Speaker 4 (m)	18 (67%)	9 (33%)	27
Speaker 5 (m)	25 (83%)	5 (17%)	30
Speaker 6 (f)	19 (68%)	9 (32%)	28
All Speakers	141 (70%)	60 (30%)	201

3.4.1. *The influence of prosodic boundary, accent and vowel on [k]*

The first factor examined in this section is the prosodic boundary following [ŋ] and [ŋk]. The prosodic domains assumed in this paper are rather traditional and have been proposed by Nespor and Vogel (1986) and Selkirk (1984). The prosodic hierarchy for higher levels is reproduced in (5).

(5) Prosodic hierarchy

Intonation Phrase	(IP)
Phonological Phrase	(PhP)
Prosodic Word	(PW)
Foot	(F)

Both the speaker and the hearer have to rely on indicators of phrasing in order to parse the utterances in prosodic phrases (see for instance the experimental results of Schafer 1997 for English and Bader 1998 or Muckel 2001 for German). A certain number of cues can be used to this aim, like duration and tones in the form of breaks and boundary tones. Segmental alternations also signal prosodic boundaries. Glottal stops, aspiration and final devoicing are well-known phenomena in German for the delimitation of the foot and higher levels (see Féry 1995, 2003), and the alternation between presence and absence of a stop after a dorsal nasal can be added to this list. The probability of finding a devoiced obstruent, a glottal stop and so on, increases with the size of the prosodic domain. The higher the prosodic domain, the more probable is the presence of such cues. It can be speculated that the release accompanying a voiceless stop is an additional cue to finality of a prosodic domain. On the contrary, the articulation of a nasal, a following stop plus possibly the first consonant of the following word in the middle of a phonological phrase involves the coordination of several articulators which have to be positioned rapidly after each other, without being separated by vowels. This may take too much time to be consistently and reliably executed. A realization without stop is thus explained by articulatory efficiency and hearer-oriented speech.

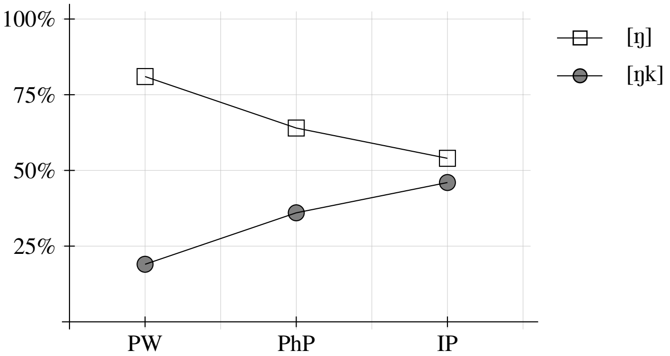
As far as the data examined here are concerned, decisions were taken as how to assign prosodic domains to utterances. In the data examined in this paper, the prosodic words were straightforward, since all realizations of the dorsal nasal considered for the results were realized at the boundary of a word which can be unambiguously analyzed as a prosodic word. The phonological phrases were decided on the basis of the syntactic structure: the right edge of a major phrase was assigned a boundary of a phonological phrase. The intonation phrases were uncontroversial since they are the locus of boundary tones and clear breaks between utterances (Pierrehumbert 1980, Pierrehumbert and Hirschberg 1990). Moreover boundaries of sentences fall together with boundaries of IPs.

A clear correlation between the kind of prosodic boundary and the presence of a stop could be established. The stronger the prosodic boundary, the more probable the occurrence of a voiceless stop. The proportion of realizations with a voiceless stop increases from 19% for a prosodic word (PW) boundary to 36% for a Phonological Phrase (PhP) boundary and 46% for an Intonation Phrase boundary (IP), as shown in (6) and (7).

(6) Prosodic domains

	[ŋ]	[ŋk]	Total
PW	87 (81%)	21 (19%)	108
PhP	25 (64%)	14 (36%)	39
IP	29 (54%)	25 (46%)	54

(7) The distribution of [ŋk] and [ŋ] according to prosodic domains



The next factor investigated is the role of a pitch accent. We postulated that being accented increases the probability to find [k], since a voiceless stop enhances the perceptivity of the word it ends.[5] And indeed, we found a confirmation of this hypothesis in our data, as shown in (8). A large proportion of target words were accented (184 words as opposed to 17 unaccented), reflecting the fact that in free speech, speakers tend to realize more accents than strictly necessary for information structural needs. All words with [ŋk] were accented, whereas 12% of the words ending with [ŋ] were unaccented.

(8) Distribution of [ŋ] and [ŋk] according to accent

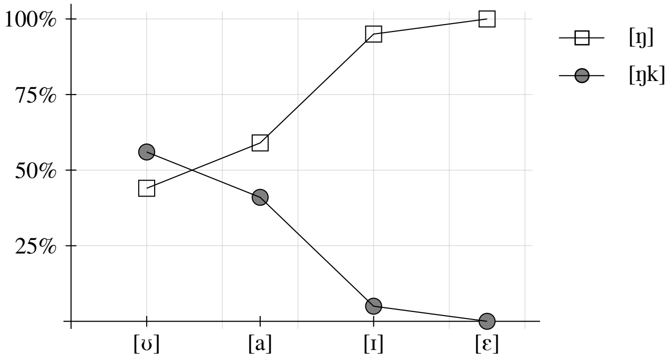
Stressed	Stressed	Unstressed
[ŋ]	124 (88%)	17 (12%)
[ŋk]	60 (100%)	0

The third factor playing a crucial role in the distribution of the allophones is the quality of the vowel immediately preceding the dorsal nasal. Compare table (9) which partitions the occurrences of [ŋ] and [ŋk] according to the preceding vowel, [ʊ], [a], [ɪ] and [ɛ]. [k] emerged much more often in the corpus after a back vowel than after a front one: 56% after [ʊ] and 41% after [a], as compare to 5% after [ɪ] and 0% after [ɛ].[6] Important is the proportion between the realizations of [ŋ] and [ŋk] after each vowel, given here in percentage, not the absolute number of occurrences since these are biased by the large number of occurrences of *Peng* and of *Wolfgang*, both names which appeared very often in all narrations see the figures in (3). Words ending in [ʊŋ] are also very frequent, but words ending in [ɪŋ] are definitely less so, as can be gathered from table (9). Figure (10) is a graphic illustration of the distribution of [ŋ] and [ŋk] after vowels.

(9) Distribution of [ŋ] and [ŋk] according to the preceding vowel

	after [ʊ]	after [a]	after [ɪ]	after [ɛ]	Total
[ŋ]	23 (44%)	43 (59%)	20 (95%)	55 (100%)	141
[ŋk]	29 (56%)	30 (41%)	1 (5%)	0	60
Total	52	73	21	55	201

(10) The influence of the vorwel



The fact that in most cases no stop is realized after a front vowel is not the result of an accidental gap due to the absence of words with a front vowel at the end of a PhP or IP. In a comparison between the number of words ending in the different vowels at each prosodic boundary, no crucial difference can be found. At the end of a PW, we find 46% of the words ending in [ʊ], 60% of those ending in [a], 67 % of those ending in [ɪ] and 47% of those ending in [ɛ]. For a PhP, we have 23%, 16%, 9% and 24% of the words with [ʊ], [a] [ɪ] and [ɛ] respectively. At the end of an IP: 31%, 23%, 24% and 29%. In other words, the proportion of words at the end of each prosodic domain does not differ dramatically for each vowel.

Table (11) shows the proportion of the allophonic variants after each vowel and for each prosodic category.

(11) Distribution according to vowels and prosodic domains

	after [ʊ]	after [a]	after [ɪ]	after [ɛ]
Word finally:				
[ŋ]	16 (66,6%)	32 (57%)	13 (93%)	26 (100%)
[ŋk]	8 (33,3%)	12 (43%)	1 (7%)	-
PhP finally:				
[ŋ]	7 (58%)	3 (25%)	2 (100%)	13 (100%)
[ŋk]	5 (42%)	9 (75%)	-	-
IP finally:				
[ŋ]	-	8 (47%)	5 (100%)	16 (100%)
[ŋk]	16 (100%)	9 (53%)	-	-
Total	**52**	**73**	**21**	**55**

An interesting result of table (11) is that at the end of an IP, there is a 100% probability to find [ʊŋk] (and 0% to find [ʊŋ]).

To sum up, we think that there is a real effect of the vowel quality on the two allophones, and that the distribution of the words ending with the different vowels at different prosodic domains in our small corpus cannot explain the difference we find.

3.4.2. *Word length, stress location, and morphemic structure: the factors which play no role in the emergence of [k]*

Turning now to the factors which do not seem to play any role for the distribution of [ŋk] and [ŋ], no correlation between the number of syllables and the presence of [k] could be established. In longer words, the lexical stress was nearly always on a nonfinal syllable (the only exceptions being *Gesang* singing' and entlang, 'along'), and, of course, monosyllabic words are stressed on their unique syllable. However, the relationship between the number of syllables and the presence of a stop will be marginal at best, since words with final [ʊ] are always multisyllabic, and always have nonfinal stress, whereas a large amount of our words ending with [ɪ] or [ɛ] were monosyllabic. Since words with final [ʊ] are more often stressed on a non-final syllable than words with final front vowel, it cannot be the case that the presence of a stress on the syllable itself is decisive. A final question arises for words with [a]. An evenly distribution of monosyllabic and finally-stressed *lang, Fang, Hang*, on the one hand, and polysyllabic and initially-stressed *Wolfgang*, on the other hand is found in all contexts, a fact leading to the conclusion that there is no correlation between number of syllables and distribution of [ŋk] and [ŋ].

It has been shown that the pattern of final *t* deletion in English is sensitive to the morphological structure of the words. Simple roots are more subject to *t*-deletion than stems which in turn are more often affected than derived words (see for instance Guy 1991). It might be interesting to find out whether a similar pattern is acting in the variation examined here.

Of the words listed in (3), some are just roots, but many are morphologically complex, like *langweilig* or *Achtung*. The majority of complex words are derived with a suffix and/or are compounds (*Fischfang, Mietwohnung*).

It is evident that, depending on the final vowel, words are distributed into different word classes. Words ending with [ʊ], for instance, are always derived with the suffix *-ung*. Words ending in [a], on the other hand, display more variation. Monomorphemic words do not seem to show any preference for the one or the other realization: *lang*, and *Wolfgang*, for instance, are found in both variants. The same is true for most complex words, like those derived with the suffix *-ung*. The only interesting generalization that seems to emerge from our limited set of data is the fact that strong verbs with ablaut (like *schwang, klang, sang* and *ging*) are always realized without stop.

The corpus used in this experiment is unfortunately much too small to be conclusive, and we leave it to further research to find out whether the morphological structure, the number of syllables and the location of lexical stress could influence the realization of the dorsal variant in one way or another.

3.5. Discussion

The result of this investigation is that a contextually conditioned distribution of the variable realizations of the dorsal nasal in the Northern German dialect could be established. Some of the variables identified deliver categorical results, and some are just tendencies. Putting together the results presented in the preceding section, the following factors play a role in the distribution of [ŋ] and [ŋk].

- First, there is an interindividual variation. Speaker 5 had the fewest occurrences of [ŋk] with 17% and speaker 3 the most with 42%.
- Second, the prosodic boundary: the stronger it is, the more probable is an occurrence of [k]. This is true for back vowels to a larger extent than for front vowels.

- Third the accent: the presence of [k] correlates with pitch accent. There were, in the corpus, not a single unaccented word with [k]. But the reverse is not true: there were accented words without [k].
- Fourth, the vowel quality: the emergence of a final [k] correlates with the preceding vowel. Only back vowels are regularly followed by a voiceless stop, [ʊ] more than [a]. A front vowel is nearly always followed by just a dorsal nasal, though the correlation is not necessary. Speakers of Northern German do sometimes realize *eng* 'tight' as [ɛŋk] or *Ding* 'thing' as [dɪŋk].

No correlation between word length and presence of [k] could be established. Other factors playing no role in the corpus were the place of the lexical stress, and the morphological category. These factors are accordingly ignored in the presentation of the model in the next section.

4. OT-CC and factor-driven gradient model

The preceding section has shown that different phonological factors influence the distribution of [ŋ] and [ŋk], and the aim of this section is to sketch a model of phonology, based on McCarthy's (2007) OT-CC, as well as on Boersma (1998) and Boersma and Hayes' (2001) Gradual Learning Algorithm, which can take these factors into consideration.

As has been discussed in section 2, traditional generative models are inappropriate for the treatment of variable and gradient data. Rules and constraints determine a unique surface form, or a unique optimal candidate, and leave no place for variation or optionality between different candidates.[7] In its standard version, Optimality Theory is as categorical as any other generative model, though constraint violability and rankability renders it a potentially more flexible framework to account for optionality. And indeed, there have been a number of proposals in the literature taking advantage of these properties in order to integrate variation into generative phonology. First, the recent Maximum Entropy model of Hayes and Wilson (2008), a model of acquisition of phonology loosely based on OT, but relying on the idea that no constraint is innate. This approach is not further considered in this section, mainly because it has appeared after completion of the present paper. Second, the multiple grammar approaches of Anttila (2001) and Reynolds and Nagy (1994), which envisage tied constraints as multiple grammars. Variation arises from the competition of distinct grammars. In Anttilas's wording:

"The number of grammars that generate a particular output is proportional to the relative frequency of this output." The main idea of such approaches is that a candidate which emerges as the optimal one in the majority of competing grammars is realized more often than a candidate only seldom selected as optimal.

Boersma's stochastic model relies on the same idea, although its formal architecture is radically different. Before introducing the factor-driven gradient model to be proposed, the main features of stochastic OT are sketched in the next subsection.

4.1. Stochastic OT: Functional OT and Gradual Learning Algorithm

Functional Grammar (Boersma 1998), as well as Boersma and Hayes' 2001 Gradual Learning Algorithm, GLA, have been specifically conceived to take care of free variation, gradient grammaticality judgments and frequency dependent time ordering in the process of acquisition. These variants of OT are similar to the standard version of OT in many ways, but instead of being based on a complete ranking of the constraints, they assume that constraints have a ranking value on a continuous scale of constraint strictness. In Functional OT, the distance between two constraints is meaningful. A shorter distance implies that the relative ranking between the two constraints is less fixed. At a certain point of time, the constraints may occupy partly the same space in the scale, a fact that Boersma and Hayes explain by constraints' overlap. This is to be understood as follows: At the time of a particular evaluation, the constraints' values are temporarily perturbed by a random factor and constraints behave as if they were associated with a range of values rather than with single points. This effect is called the 'evaluation noise.' It is an amount of normally distributed noise and is temporarily added to the ranking value of each constraint. The amount of noise is calculated independently for each evaluation.

If the ranking values of two constraints, A $\gg$ B, are close, the ranking of the constraints at evaluation time will sometimes be the reverse of their "normal" ranking, depending on the accidental values of the noise component. In other words, in a certain percentage of the evaluations, B will outrank A although it is ranked somewhat lower, and, when this happens, the second best candidate wins. At a given moment (as for instance in language acquisition), the noise value of the constraints which are further apart can be so large that even a constraint which is usually dominated, can temporarily become dominant. If the ranges of two constraints overlap, there will be free variation. Constraint ranges are thus interpreted as probability distributions, and account in this way for noisy events. A normal distribution has a single peak in the center, with values around the center being most probable, and declining toward zero on each side. Values are less probable the further they are from the center.

As an extension of Functional Phonology, Boersma and Hayes (2001) develop the Gradual Learning Algorithm, an OT algorithm for learning constraint ranking. They illustrate its working with examples of free variation in the distribution of glottal stop and glides in Ilokano, of output frequency in the Finnish genitive and of gradient well-formedness judgments in the distribution of light and dark [l] in English. The Gradient Learning Algorithm needs two kinds of inputs, frequencies of distribution and OT constraints. The output of the algorithm is a ranking of the constraints on a continuous scale, as explained above. The GLA calculates the strictness of the constraints relatively to each other. The ranking is expressed in terms of values, which can thus be closer or further apart.

The data examined in this paper can be accounted for by a stochastic model, as will be shown in section 4.3, but the allophony between [ŋ] and [ŋk] is not free. Rather it is influenced by different factors, a fact which introduces a complication for the model.

4.2. OT-CC and the allophony between [ŋ] and [ŋk]

Before we can discuss the model of variation in the next subsection, it is important to introduce the constraints necessary to account for the allophony between [ŋ] and [ŋk]. An additional difficulty related to this allophony is that one of its allophones, namely [ŋ], results from an opaque process (see also van Oostendorp, this volume for opacity in OT).

Assuming first that [ŋ] is always the product of an assimilation of a coronal nasal to a following dorsal stop, a constraint AGREE(dorsal) in (12a) is needed. In German monomorphemes, this constraint is inviolable. To illustrate the effect of this constraint, a word like *Bank*, which preserves the final [k], is used in (13). A constraint against deletion of [k], MAX[k], formulated in (12b), is also needed. A third constraint, IDENT(coronal), formulated in (12c), which requires that an input feature [coronal] is preserved in the output, is violated by the grammatical candidate. Tableau (13) illustrates the effect of (12) with a non-alternating word.

(12)

a. AGREE(dorsal): Adjacent nasal and dorsal stop agree in dorsality.
b. MAX[k]: No deletion of the segment [k].
c. IDENT(coronal): No change of the feature [coronal].

(13) OT account of *Bank* 'bank'

/bank/	AGREE(dorsal)	MAX[k]	ID(coronal)
☞ a. baŋk			*
b. bank	*!		
c. ban		*!	

More intricate is the account of the allophony between [ŋ] and [ŋk/ŋk]. The first variant, without [g] or its finally devoiced correspondent [k], is the product of an opaque counterbleeding interaction (overapplication), since the trigger of the assimilation [g] is not present on the surface. Opaque situations are notoriously difficult to account in an optimality theoretic model, and have been the cause of many amendments to the standard OT model (see MCarthy 2003, Kiparsky 2003, Ito and Mester 2003 to cite just a few). We propose to use McCarthy's (2007) Candidate-Chain version of OT (OT-CC) for the case at hand. This extension of standard OT represent candidates as ordered chains of candidates with minimal faithfulness violations of increasing harmony.[8] OT-CC makes a crucial use of constraints responsible for precedence relations between violations, called Prec(A,B), which demand that both violations are present and that A is violated before B in the candidate chain. In the case of [ŋ], the violation of IDENT(coronal) must precede the violation of MAX[g], thus assimilation must take place before deletion, see (14b). The trigger for deletion is the constraint *[ŋg], as in (14a), which has been proposed by several phonologists, in this special variant (see Ito and Mester 2003) or in a more general variant prohibiting a sequence of a nasal plus voiceless obstruent (Pater 1996). The effect of these constraints is illustrated in Tableau (15).

(14)

a. *[ŋg]: A sequence of a dorsal nasal plus stop is prohibited.
b. Prec(ID(cor),MAX(g)): In a candidate chain, the violation of IDENT(coronal) must precede the violation of MAX[g].

Candidate a. is the winner because it fulfills not only *[ŋg] which is violated by candidate b., but also the Prec constraint, which is violated by the transparent candidate c.

(15) OT-CC account of the [ŋ] variant

/ding/	Agr (dors)	*[ŋg]	Prec(Id(cor), Max(g))	Max[g]	Id(cor)
☞ a. diŋ <ding, diŋg, diŋ>				*	*
b. diŋg <ding, diŋg>		*!	*		*
c. din <ding, din>			*!	*	

If input /g/ is realized, as it is in the allophone [ŋk], it has no choice but to become a voiceless stop because of Final Devoicing (FD), as formulated in (16a). This constraint must be ranked higher than both Max[g] and Ident(voice), in (16b-c), see Tableau (17).

The second allophone [ŋk] is not opaque and thus does not require a Prec constraint. However, since it is in an allophonic relation with [ŋ], its evaluation interacts with the evaluation shown in (15). We propose a second Prec constraint, namely (16d), which orders a violation of Ident(coronal) before a violation of Ident(voice), see candidate a. in Tableau (17). This Prec constraint requires a different chain of candidates, namely one in which final devoicing is ordered directly after assimilation. A violation of Ident(voice) bleeds a violation of Max(g) since if /g/ is devoiced to [k], it is no longer susceptible to be deleted under the influence of *[ŋg].

(16)

a. FinalDevoicing: A foot-final obstruent is voiceless.
b. Max[g]: No deletion of the segment [g].
c. Ident(voice): No change of the feature [voice].
d. Prec(Id(cor),Id(voi)): In a candidate chain, the violation of Ident(coronal) must precede the violation of Ident(voice).

Tableau (17) illustrates the choice of *diŋk*. It contains both the constraints necessary for *diŋ*, which were shown in Tableau (15), and the new constraints. Agr[dorsal] and *[ŋg] are not shown here for reasons of space.

(17) OT account of the [ŋk] variant

/ding/	FD	Prec(Id(cor), Id(voi))	Prec(Id(cor), Max[g])	Max(g)	Id(voi)	Id(cor)
☞ a. diŋk <ding, diŋg, diŋk>			*		*	*
b. diŋ <ding, diŋg, diŋ>		*!		*		*
c. diŋg <ding, diŋg>	*!	*	*			

Candidate b. is eliminated because it does not violate Id(voice), as required by Prec(Id(cor),Id(voi)). Candidate c. violates also FinalDevoicing. We do not go back to the table (15), but it should be clear that the two Prec constraints are ranked in the opposite direction in (15) and in (17). In other words, the allophony is accounted for by different ordering of the two Prec constraints. (18) shows the ranking of the constraints delivering the allophone [ŋ], and (19), where the Prec constraints come in the reverse order, the one delivering [ŋk].

(18) Ranking for [ŋ]
FD, Agree(dorsal), Max[k], *[ŋg] ≫ Prec(Id(cor),Max(g)) ≫
Prec(Id(cor), Id(voi)) ≫ Max[g], Ident(coronal), Ident(voice)

(19) Ranking for [ŋk]
FD, Agree(dorsal), Max[k], *[ŋg] ≫ Prec(Id(cor), Id(voi)) ≫
Prec(Id(cor), Max(g)) ≫ Max[g], Ident(coronal), Ident(voice)

What determines the choice between (18) and (19) in Northern German is the subject of next section. We propose that the choice is determined by internal phonological factors, rather than by random noise.

4.3. Factor-driven gradient model (OT-FDG)

The preceding section has shown that it is the ranking of the two Precedence constraints that determines which of the two allophones are chosen in a particular evaluation. But it remains to be demonstrated how the factors bearing on the allophony and introduced in section 3, are formalized into the OT grammar, in order to influence the choice of one or the other candidate.

We do not propose a fully-fledged model of gradience in this paper, since this is beyond the scope of this paper. Instead, we sketch how the integration of different factors acting on the distribution of [ŋ] and [ŋk] can be part of the grammar itself. We propose an extension of Functional Phonology, based on the influence of each factor on the probability of the output, that we call factor-driven gradient model (FDG). The difference between GLA and FDG is that the latter model does not need the exact distribution of the allophones in all instances, but rather the probability of each allophone is calculated by the grammar itself. It relies on a baseline and a knowledge of the effect that each of the factors exerts on the probability to get one or the other variant.

The initial situation, shown in (20), is the default distribution: the total amount of [ŋk] and [ŋ], as given in table (4). In 70% of the cases, Prec(Id(cor), Max(g)) dominates Prec(Id(cor),Id(voi)), as (18) and in 30% of the cases, the reverse ranking is active, as in (19). The two Prec constraints have a variable ranking which is best accounted for in a stochastic version of OT, with overlapping constraints. This overall distribution of the two variants is considered here as the baseline.

(20)

ŋ	ŋk
70% (ranking (18))	30% (ranking (19))

However, as was shown in section 3, the preference for one or the other variant can be increased depending on several context-dependent factors, like prosodic boundary strength, pitch accent and vowel quality. In other words, the optionality is not random or created by momentary noise, but rather it is triggered and motivated by factors which have a direct influence on the grammar. Specific constraints influence the ranking of the Prec constraints.

Considering in a first step the prosodic domain boundary, it was shown that a strong prosodic boundary increases the probability of an occurrence of [k]. Both the end of a PhP and of an IP have a positive effect on [k], and a PW boundary has a negative effect. A markedness hierarchy (Prince and Smolensky 1993-2004) can be assumed, that correlates the realization of a final stop with the end of different prosodic domains, as in (21b-e). If there is no stop (or, in our case, if the allophone [ŋ] is chosen), the constraint is violated. As usual for such hierarchies, they express universal generalizations, and are supposed to be ranked in the same order in all languages. The ALIGN constraints in (21) explicitly ask for a realized stop at the end of prosodic constituents, and imply that, if a stop is realized optionally, it will be found more often at the end of a higher prosodic domain than at the end of a lower one. In German, this constraint hierarchy is relevant only for the sequence /ng/, since this is the only consonant sequence which can be optionally realized with or without a stop (see Féry 2003). All other consonants are obligatorily realized, so that the hierarchy is vacuous for them: they are protected by higher-ranking faithfulness constraints. In the ranking (21a), DEP, the constraint prohibiting epenthesis, is ranked higher than the alignment constraints, and this hierarchy has the effect that no stop will ever be inserted just to satisfy them.

(21) Aligment constraints
  a. DEP $\gg$ ALIGN(IP, Stop, R) $\gg$ ALIGN(PhP, Stop, R) $\gg$ ALIGN (PW, Stop, R)
  b. DEP: No epenthesis.
  c. ALIGN(IP, Stop, R): The right boundary of an IP is aligned with a stop.
  d. ALIGN(PhP, Stop, R): The right boundary of a PhP is aligned with a stop.
  e. ALIGN(PW, Stop, R): The right boundary of a PW is aligned with a stop.

Since the ALIGN constraints have the effect to favor the presence of a stop in the relevant environment, we can think of the Alignment constraints as changing the proportion of [ŋ] vs. [ŋk]. In other words, the grammar of the language as a whole delivers a certain ranking value of two Prec constraints and some additional constraints have a contextual effect which changes their ranking in a given proportion.

As an illustration, consider first the end of an IP. In this case, the probability of finding [k] is 48%. This constraint increases the probability of [k] as compared to the default distribution shown in (22) by 18%.

(22)

ŋ	ŋk
52% (ranking (18))	48% (ranking (19))

In terms of ranking of constraints, the two Prec constraints are still ranked in the baseline order shown in (20), but ALIGN(IP, Stop, R) is ranked between them, and overlap with both of them. Since it requires a stop, it increases the probability of a realized [k] by a certain factor which is proportional to the amount of overlapping, and which can be calculate from the exact digits given in section 3.[9] In other words, two constraints are now ganging up for the realization of a final stop.

The effect of a PhP boundary is slightly different, since it increases the probability of realizing [k] by 8% only, relatively to the default 30%. Like ALIGN(IP, Stop, R), it is ranked between the two Prec Constraints, but overlap less with the first one than ALIGN(IP, Stop, R) does. And as far as ALIGN(PrW, Stop, Right) is concerned, it is active when only a Prosodic Word boundary is present after the word containing the dorsal nasal, that is when the word does not end a higher prosodic domain at the same time. In this particular context, the influence on the presence of [k] is negative.

The effect of specific preceding vowels has been quantified as well, and can be treated in the same way as the alignment constraints just examined. In the case of the vowel, too, the context is exactly defined. First, consider [ʊ]. In order to integrate the effect of this vowel on the distribution of [ŋk], we need a constraint to the effect that a back high vowel will be followed by the sequence [ŋk], which can be tentatively considered as a kind of backness harmony effect. A formulation is proposed in (23).

(23) BLACKHIGHVOWEL+[ŋ]+BACKOBSTRUENT (*ʊŋk):
A back vowel [ʊ] requires the realization of all back consonants.

According to table 17, 60% of words with [ʊ] are realized with [k], which means that this constraint has an even more dramatic effect on the presence of [k] than the IP boundary. The distribution of [ŋ] and [ŋk] because of [ʊ] is shown in (24).

(24)

ŋ	ŋk
40% (ranking (18))	60% (ranking (19))

If *ʊŋk is treated in the same way as ALIGN(IP, Stop, R), that is, it is ranked between the two Prec constraints, it overlaps them both and it increases the probability of realization of [k]. In some situations, the effects of ALIGN(IP, Stop, R) and of *ʊŋk will be added to each other, and a ganging up effect through overlapping will be obtained.

The corpus did not contain occurrences of [k] after a mid front vowel [ɛ], thus the effect of this vowel is categorical (as long as the corpus discussed above is the only source of distributional data). In association with such a vowel, only the variant [ŋ] is found. The constraint (25) might be undominated.

(25) *FrontMidVowel+[ŋ] +BackObstruent (*ɛŋk):
A front vowel [ɛ] blocks the realization of k after [ŋ].

To sum up, the factors that were shown to influence one or the other variant of the dorsal nasal, are expressed in the form of individual constraints. Their effect is to increase or decrease the probability of the two variants. Like the GLA, FDG does not deliver concrete outputs, but just probabilities for variable outputs. The difference between the two models is to be found in the controlled influence of the factors, as opposed to the effect of noise in GLA. The chance to utter [ŋk] or just [ŋ] is due to the impact of the individual factors. As in GLA, all constraints participating in an allophonic alternation are ranked in a fixed order, but in FDG, constraints come with a factor influencing the distance between two crucially conflicting constraints. Factors associated with different constraints can also be cumulative, and become so strong that the ranking of the crucial constraints is reversed, or as illustrated with (25), an undominated constraint can overrule the effect of the conflicting constraints. Two or more constraints, each with their own factor, can interact and increase the effect in an additional way. This happens if a word ending in /ʊng/ like *Zeitung* is at the end of an IP and accented. In this case, the probability to realize [ŋk] reaches 100%.

This section has just sketched some elements of an OT model based on concrete factors influencing allophony in a gradient way. It is far from being worked out, but we hope that it provides some elements for integrating internal factors of variation into the grammatical model itself.

5. Conclusion

It has been repeatedly observed that the Standard variety of German and its Northern variant differ in their realization of the word-final sequence [ŋg]. In the Standard variety of German, the final stop has been claimed to be deleted, resulting in opaque [ŋ], whereas it is assumed to be subject to Final Devoicing in Northern German, resulting in [ŋk]. This difference has been accounted for in categorical versions of generative phonology (see Wurzel 1980, Hall 1992, Ramers and Vater 1992, Yu 1992, Wiese 1996, Ito and Mester 2003 and Féry 2003 among others).

An OT analysis of the allophony between opaque [ŋ] and transparent [ŋk] is couched in McCarthy's new version of Optimality Theory, CC-OT, which contains elements of derivational phonology in the obligatory ordering of constraints. The two allophones require two different ordering of Precedence Constraints. This model can be used for the cross-dialectal difference, but also for some gradient data in a single dialect.

A detailed analysis of a corpus of spontaneous speech of speakers from the Berlin-Brandenbourg region (Northern German) reveals that the realization of this sequence is far from being categorical in the way assumed in the past. Both [ŋ] and [ŋk] are frequent in this dialect, though in different proportions. If all contexts are taken together, [ŋ] is more frequent than [ŋk], in a proportion of 70 to 30%. Additionally, the proportion of occurrences of the allophones is influenced by the contexts in which the segment sequences occur. A word ending in a dorsal nasal is often pronounced with a final voiceless stop [k] when it ends a higher prosodic boundary, such as an Intonation Phrase. An accented word is also more likely to be realized with a final stop, and the same is true for words whose last vowel is a back vowel.

An increasing amount of researchers (as for instance Pierrehumbert 2003, Bybee 2001, Cohn 2006, Frisch 2000, Anttila 2001, Boersma 1998, Boersma and Hayes 2001, Hayes and Wilson 2008, Labov 1994, Steriade 2001 and many others) have emphasized the need to integrate gradience and variation into formal generative phonology. Traditional generative models of phonology cannot account for the kind of conditioned variability found in the realization of /ng/. In this paper, it is proposed that stochastic Gradual Learning Algorithm (GLA) (Boersma 1998, Boersma and Hayes 2001), an OT model which has explicitly been conceived to account for gradient data, is a first step to account for our data. We sketch some ideas for an extension of this model, called factor driven gradient model (FDG). Instead of relying on random noise to explain temporarily optionality between two allophones, as in GLA, in FDG the factors bearing on the distribution of the data are part of the phonology formalism. Overlapping context-independent constraints define a baseline for the gradience, but the probability for the emergence of one or the other candidate is influenced by context-sensitive constraints.

Notes

1. This paper is part of the project A1 of the Research Group "Conflicting Rules" of the University of Potsdam. Some of the results have been presented at the Meertens Institute in Amsterdam in September 2003 and at the DGfS meeting in Mayence in February 2004. Many thanks to Birgit Alber, Marc van Oostendorp, Tonio Green, Frank Kügler and Frans Hinskens for helpful comments and discussions. Thanks to Daniela Berger and Anja Arnhold for technical support. The experiment was designed and executed by the second and third authors, in the framework of a seminar on variation in phonology directed by the first one. The remaining of the work is the sole responsibility of the first author.
2. In this paper, only environments are considered which can be the locus of g-deletion, that is, syllable-final or onsets of non-moraic syllables. Cases in which [g] is pronounced because it is followed by a full vowel, as in *Tango, Pinguin or diphthongieren* 'to diphthongize' are ignored (see Féry 2003 for an OT account of obligatory [g] in terms of onset).
3. Words in which final [ŋk] is followed by further consonants, like in *singst* or *Angst*, as well as words with a medial ambisyllabic dorsal nasal followed by a schwa syllable, as in *singen, Klingel, engen, Wohnungen* and *gedrungene*, were not considered in the results and discarded altogether. In these words, the obstruent is never realized.
4. Some statistical test were performed to calculate the significance of the results, as for instance Wilcoxon-test and chi-square (thanks to Wang Bei and Robin Hörnig for advices), but due to the small amount of data, no really interesting statistical claims could be made, except maybe for the following one: for all speakers, the difference between [ŋ] and [ŋk] was significant.
5. Birgit Alber [p.c.] proposes that [k] is realized more often when the word is accented because accenting prevents fast speech phenomena (i.e. deletion) from happening.
6. In German, [a] patterns phonologically with the back vowels, as can be seen from different alternations, like the distribution of the dorsal fricative (*ach*-Laut and *ich*-Laut) and the umlaut which fronts back vowels, including [a].
7. There have been some proposals to integrate variability into the transformational format, as testified by Labov's so-called "variable rule" (1969), and mathematical modeling of it (Cedergren and Sankoff 1974, Sankoff and Labov 1979, see also Fasold 1991). Variable rules have been mostly used to integrate 'external' sociological variables into phonological rules. In the case at hand, the intention is different, namely the elaboration of a model of 'internal' phonological effects influencing the phonology.
8. In McCarthy (2007), the candidates are 4-tuples consisting of {input, output, the violations with their locus and the ordering between them} The version we present here is largely simplified.

9. We do not try to develop an exact mathematical model of FDG, but restrict ourselves to the outlines of the analysis.

References

Anttila, Arto
2001 Variation and Phonological Theory. In *The handbook of language variation and change,* J. K. Chambers, P.Trudgill and N. Schilling-Estes (eds.), 206-243. Oxford: Blackwell.

Bader, Markus
1998 Prosodic influences on reading syntactically ambiguous sentences. In *Reanalysis in sentence Processing,* Janet D. Fodor and F. Ferreira (eds.), 1-46. Dordrecht: Kluwer.

Boersma, Paul
1998 *Functional Phonology. Formalizing the interactions between articulatory and perceptualdrives.* Ph. D. thesis. University of Amsterdam. The Hague: Holland Academic Graphics.

Boersma, Paul and Bruce Hayes
2001 Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32(1): 45-86.

Bybee, Joan
2000a The phonology of the lexicon. Evidence from lexical diffusion. In *Usage-based models of language,* M. Barlow and S. Kemmer (eds.), 65-85. Stanford: CSLI Publications.

Bybee, Joan
2000b Lexicalization of sound change and alternating environments. In *Laboratory V: Language Acquisition and the Lexicon,* Michael Broe and Janet Pierrehumbert (eds.), 250-268. Cambridge: Cambridge University Press.

Bybee, Joan
2001 *Phonology and Language Use.* Cambridge: Cambridge University Press.

Bybee, Joan
2003 Mechanisms of Change in Grammaticization: The Role of Frequency. In *Handbook of Historical Linguistics,* Richard Janda and Brian Joseph (eds.), 602-623. Oxford: Blackwell.

Cedergren, H.J. and D. Sankoff
1974 Variable rules: Performance as a statistical reflection of competence. *Language* 50: 333-355.

Cohn, Abigail
2006 In there gradient Phonology? In *Gradience in Grammar,* Gisbert

Fanselow, Caroline Féry, Ralf Vogel and Matthias Schlesewsky (eds.), 25-44. Oxford: Oxford University Press.

Darcy, Isabelle, Franck Ramus, Anne Christophe, Katherine Kinzler and Emmanuel Dupoux
this volume Phonological knowledge in compensation for native and non-native assimilation.

Fasold, R.
1991 The quiet demise of variable rules. *American Speech* 66: 3-21.

Féry, Caroline
1995 Alignment, syllable and metrical structure in German. Habilitationsschrift. *SfS-Report* 02-95. Tübingen: University Tübingen.

Féry, Caroline
2003 Nonmoraic syllables in German. In *The Syllable in Optimality Theory,* Caroline Féry and Ruben van de Vijver (eds.), 213-237. Cambridge: Cambridge University Press.

Frisch, Stefan A.
2000 Temporally Organized Lexical Representations As Phonological Units. In *Acquisition and the Lexicon: Papers in Laboratory Phonology V,* Michael Broe and Janet Pierrehumbert (eds.), 283-298. Cambridge: Cambridge University Press.

Guy, Gregory R.
1991 Explanation in variable phonology. *Language Variation and Change* 3: 1-22.

Hayes, Bruce and Colin Wilson
2008 A Maximum Entropy Model of Phonotactics and Phonotactic Learning. *Linguistic Inquiry* 39(3): 379-440.

Hall, Tracy Alan
1992 *Syllable Structure and Syllable-Related Processes in German.* Tübingen: Niemeyer.

Isačenko, Alexander
1963 Der phonologische Status des velaren Nasals im Deutschen. *Zeitschrift für Phonetik* 16: 77-84. Reprint in *Vorschläge zu einer strukturalen Grammatik des Deutschen,* Hugo Steger (1970) (ed.), 468-479 Darmstadt: Wissenschaftliche Buchgesellschaft.

Ito, Junko and Armin Mester
2003 On the sources of opacity in OT: Coda processes in German. In *The Syllable in Optimality Theory*, Caroline Féry and Ruben van de Vijver (eds.), 271-303. Cambridge: Cambridge University Press.

Jacewicz, Ewa, Joseph Salmons and Robert Allen Fox
this volume Prosodic conditioning, vowel dynamics and sound change.

Keating, Patricia A.
1987 A Survey of Phonological Features. *UCLA Working Papers in Phonetics* 6: 124-150. Los Angeles, California.

Kiparsky, Paul
2000 Analogy as Optimization. In *Analogy, Levelling, Markedness. Principles of Change in Phonology and Morphology. Trends in Linguistics. Studies and Monographs 127,* Aditi Lahiri (ed.). Mouton de Gruyter.

Kiparsky, Paul
2003 Syllables and Moras in Arabic. In *The Syllable in Optimality Theory,* Caroline Féry and Ruben van de Vijver (eds.), 147-182. Cambridge: Cambridge University Press.

Kirchner, Robert
2001 *An Effort-Based Approach to Consonant Lenition*. Ph. D. thesis, University of California, Los Angeles. New York, NY: Routledge.

Kroch, Anthony
2001 Syntactic change. In *The Handbook of contemporary syntactic theory,* M. Baltin and C. Collins (eds.), 699-729. Oxford: Blackwell Publishers.

Labov, William
1966 *The Social Stratification of English in New York City*. Washington, DC: Center for applied linguistics.

Labov, William
1994 *Principles of Linguistic Change. Internal factors.* Oxford: Blackwell.

McCarthy, John J.
2003 Sympathy, Cumulativity, and the Duke-of-York Gambit. In In *The Syllable in Optimality Theory,* Caroline Féry and Ruben van de Vijver (eds.), 23-76. Cambridge: CUP.

McCarthy, John J.
2007 *Hidden Generalizations: Phonological Opacity in Optimality Theory*. Equinox. Oxford.

Muckel, Sandra
2001 *Wortstellungseffekte beim Satzverstehen: Zur Rolle syntaktischer, verbspezifischer und prosodischer Informationen*. Ph. D. thesis. Leipzig.

Nespor, Marina and Irene Vogel
1986 *Prosodic phonology.* Dordrecht: Foris.

Ohala, John J. and Manjari Ohala
1995 Speech perception and lexical representation: The role of vowel nasalizaiton in Hindi and English. In *Papers on Laboratory Phonology IV: Gesture, Segment, Prosody.* Bruce Cornell and Amalia Arvaniti (eds.), 41-60. Cambridge: Cambridge University Press.

Pater, Joe
1996 *NC. In *Proceedings of the North East Linguistic Society* 26, K. Kusumoto (ed.), 227-239.

Pierrehumbert, Janet
1980 *The phonology and phonetics of English intonation.* Ph. D. thesis. MIT. Cambridge, Mass.

Pierrehumbert, Janet
2003 Probabilistic Phonology: Discrimination and Robustness. In *Probabilistic Linguistics,* Rens Bod, Jennifer Hay and Stefanie Jannedy (eds.), 177-228. Cambridge, MA: MIT Press.

Pierrehumbert, Janet, Beckman, Mary E. and Robert Ladd
2000 Conceptual Foundations in Phonology as a Laboratory Science. In *Phonological Knowledge: Conceptual and Empirical Issues,* N. Burton-Roberts, P. Carr and Gerhard Docherty (eds.), 273-304. New York: Oxford University Press.

Pierrehumbert, Janet and Julia Hirschberg
1990 The meaning of intonational contours in the interpretation of discourse. In *Intentions in communications,* P. Cohen, , J. Morgan and M. Pollock (eds.), 271-311. Cambridge: MIT Press.

Prince, Alan and Paul Smolensky
1993 *Optimality Theory: Constraint Interaction in Generative Grammar.* Malden: 2004 MA and Oxford: Blackwell.

Ramers, Karl H. and Heinz Vater
1992 *Einführung in die Phonologie.* Köln: Gabel (= Klage 16).

Reynolds, B. and N. Nagy
1994 Phonological variation in Faetar: An Optimality Account. In *Chicago Linguistic Society 30-II: Papers from the Parasession on Variation and Linguistic Theory,* 277-292.

Sankoff, D. and William Labov
1979 On the uses of variable rules. *Language in Society* 3: 189-222.

Schafer, Amy
1997 *Prosodic parsing: The role of prosody in sentence comprehension.* Ph. D. thesis, Amherst, MA

Selkirk, Elisabeth O.
1984 *Phonology and syntax. The relation between sound and structure.* Cambridge: MIT Press.

Šimáčková, Šárka
this volume Variable quality of the Czech lateral liquid: A perception experiment with young Czech listeners.

Steriade, Donca
2001 Directional Asymmetries in Place Assimilation: A Perceptual Account.

In *The Role of Speech Perception in Phonology,* Elizabeth Hume and Keith Johnson (eds.), 219-250. New York: Academic Press,

Wiese, Richard
1996 *The phonology of German.* Oxford: Clarendon Press.

Wurzel, Wolfgang Ullrich
1980 Phonologie. In *Grundzüge einer deutschen Grammatik,* Karl E. Heidolph, Walter Flämig, Motsch, Wolfgang et al. (eds.) Berlin: Akademie-Verlag.

Yu, Si-taek
1992 *Unterspezifikation in der Phonologie des Deutschen.* Tübingen: Niemeyer (= *Linguistische Arbeiten* 274).

Lexical exceptionality in Florentine Italian *troncamento*

Judith Meinschaefer

1. Introduction

Influential theories of human language, like structuralism and generative grammar, have long concentrated on linguistic phenomena that are regular and categorical. While the integration of exceptions into formal models of linguistic competence has been under discussion since the beginnings of generative phonology (e.g., Chomsky and Halle 1968, Schane 1973, Zonneveld 1978), in recent years various mechanisms have been proposed within Optimality Theory (Prince and Smolensky [1994] 2004) that aim explicitly at representing both optionality and exceptionality not merely as uninteresting phenomena of linguistic performance but rather as integral components of linguistic competence (e.g., by assuming cophonologies, cf. Itô and Mester 1995, Inkelas, Orgun and Zoll 1997 for discussion, Inkelas 1998; indexed constraints, cf. Pater 2000; or multiple inputs, cf. Kager 1996). In this vein, the present paper develops a competence-oriented model of lexical exceptionality for Italian *troncamento*, a process of syncope whereby a word-final mid vowel (/e/ or /o/) is deleted when preceded by a sonorant, as in *andare* > *andar* 'go', *mangiano* > *mangian* 'they eat', *maggiore* > *maggior* 'greater'. The variety of Standard Italian under investigation is Florentine Italian, a variety in which final vowel deletion is particularly productive (Rosati 2001). The analysis is based mainly on corpus data drawn from a corpus of spontaneous spoken language (C-Oral-Rom, Cresti and Moneglia, eds., 2005).

While descriptive grammars often treat *troncamento* as an instance of largely optional phonological variation, a closer look reveals that Italian *troncamento* does present many phonological regularities, in that its application is sometimes quite obligatory and sometimes entirely blocked. Hence, the often observed optionality requires a model of phonological variation as arising from linguistic competence, not from performance alone. A further complication is that while *troncamento* applies freely to some lexemes, it appears to be blocked for many others. What is more, while the rule often applies to vowels within suffixes, its application is licensed or blocked not for particular suffixes, but for particular combinations of stem and suffix. Hence, whether *troncamento* applies or not cannot be specified for suffixes or for lexemes, but only with reference to specific word forms.

It is these patterns of lexical exceptionality in application of *troncamento* that are the central concern of this study. Following previous work by Nespor (1990), it will be argued that exceptionality in Italian *troncamento* presents a case of alternation between lexically listed forms, whose distribution is prosodically conditioned (where prosody is understood along the lines of Selkirk 1978, Nespor and Vogel 1986). Crucially, *troncamento* presents a case of alternation between word forms, not between stems or between suffixes. In the optimality-theoretic analysis developed here, the alternation is thus represented as selection from multiple inputs, where alternant distribution is accounted for by constraints on the ouptput (Mascaró 1996, Kager 1996, Perlmutter 1998). Since *troncamento* is a prosodic alternation, bound to the phonological phrase, as argued below, the listed forms to select from must be prosodic entities. In particular, it will be argued that the *troncamento* alternation involves selection from multiple prosodic words associated to a single morphological word. As to the constraints responsible for the selection from the list of alternants, it will be argued that Italian *troncamento* is rhythmically conditioned, and that selection of the truncated alternant is achieved by the interaction of a constraint against vowel lengthening and a constraint against extrametrical syllables. In contrast, the non-truncated alternant is selected at phonological phrase boundaries in order to assure that in this context the edges of the morphological word and of the prosodic word coincide. The model is developed as follows. Section 2 gives a descriptive account of Italian *troncamento*, and the phonological regularities of Italian *troncamento* are laid out in section 3. The patterns of lexical exceptionality in application of *troncamento* will be studied and formalized in section 4.

2. The phenomenon

In Standard Italian, and in particular in the Tuscan variety (Rosati 2001, cf. also Serianni 1988: 32), word-final vowels may be deleted in certain contexts. The vowel deletion rule, generally called *troncamento* in descriptive grammars, applies only to vowels which are preceded by a sonorant consonant (i.e., nasal or liquid), which itself must be preceded by a vowel. Deletion of the word-final mid vowels /e/ and /o/ is relatively frequent, but /a/ and /i/

also undergo deletion in some cases (cf., e.g., Serianni 1988: 29-33). Examples are given in (1) and (2). All examples presented in this paper are from the C-Oral-Rom corpus[1] (Cresti and Moneglia, eds., 2005), and positive as well as negative generalizations have been confirmed by two native informants.

(1)

V, infinitive *andare* > *andar* ‘go’

se	*si*	*va*	*a*	*piedi*	*bisogna*	*andar*	*presto*[2]
if	one	goes	to	foot	one.needs	to.go	soon

‘If we go by foot, we have to go soon.’

(2)

V, 3 sg pres *mangiano* > *mangian*, ‘eat’

tutti	*quegli*	*americani*	*mangian*	*verdure*
all	these	Americans	eat	vegetables

‘All these Americans eat vegetables.’

Since deletion of /a/ and /i/ applies to very few lexemes, it will not be considered here. Only unstressed final vowels undergo deletion. The lexemes whose vowels can be deleted belong to various lexical and functional categories. *Troncamento* is most regular for verbs, see (1) and (2) above, but it applies to nouns, see (3), adjectives (4), and adverbs (5), as well.

(3)

professore > *professor* ‘professor’, N, m sg

ringrazio	*e*	*saluto*	*ancora*	*il*	*professor*	*Bettini*
I.thank	and	greet	again	the	professor	Bettini

‘I thank and greet professor Bettini once again.’

(4)

migliore > *miglior* ‘better’, Adj, m sg

lo	*sport*	*è*	*la*	*miglior*	*chiave*	*per*	*aprire*	*le porte chiuse*
the	sport	is	the	best	key	to	open	the closed doors

‘Sport is the best key to open closed doors.’

(5)

bene > *ben* ‘well’, Adv
eventi *ben* *più* *gravi*
events well more serious
‘much more serious events’

Since most Italian words end in an inflectional suffix, the vowels that undergo deletion are often suffixes or parts of suffixes. In some cases, though, they belong to the stem, as in the adverbs *bene* ‘well’ and *male* ‘badly’.[3] Interestingly, not all suffixes ending in /e/ or /o/ allow vowel deletion. More precisely, *troncamento* can apply to the word-final vowel of a word form with the morphological specifications given in (6) to (8).

(6) verbs

vowel	form	example	lexically restricted?
/e/	infinitive	*andá-re* > *anda-r* ‘go’	no
/e/	3 sg pres	*vuól-e* > *vuol* ‘want’	to some verbs
/o/	3 pl	*mángia-no* > *mangia-n* ‘eat’	to some verbs
/o/	1 pl	*abbiá-mo* > *abbia-m* ‘have’	to some verbs
/o/	1 sg	*són-o* > *son* ‘be’	to *essere* ‘be’

(7) adjectives

vowel	form	example	lexically restricted?
/e/	sg	*miglióг-e* > *miglior* ‘better’	yes
/o/	m sg	*légger-o* > *legger* ‘light’	yes

(8) nouns

vowel	form	example	lexically restricted?
/e/	m sg	*colór-e* > *color* ‘colour’	yes
/o/	m sg	*mán-o* > *man* ‘hand’	yes

All other /e/s and /o/s may not be deleted. In particular, the vowels /e/ and /o/ never undergo deletion in the forms given in (9).

(9)

a. in verbs

/e/	3 sg	in non-present tenses
/o/	3 pl	in the conditional and in the perfect
/o/	1 pl	in non-present tenses
/o/	1 sg	in non-present tenses (of *essere* ‘be’)

b. in nouns and adjectives

/e/	if it is an exponent of fem pl (Serianni 1988: 32)

Seen from a morphological point of view, the vowel which undergoes deletion may thus be part of a suffix or may be a suffix itself, as described in (10). In some cases vowels which belong to the stem may be deleted, as in (10c).

(10)

a. in verbs
  /e/ part of the infinitive ending
  /o/ part of the exponent of 3 pl
  /o/ part of the exponent of 1 pl
  /e/ if it is an exponent of 3 sg pres or 1 sg pres
b. in nouns and adjectives
  /e/ if it is an exponent of sg
  /o/ if it is an exponent of m sg
c. in adverbs
  /e/ in *bene* 'well' and *male* 'badly'

Before showing how these patterns of exceptionality in Italian *troncamento* can be represented in a model of the interface between phonology, syntax and the lexicon, I shall examine in more detail the syntactically defined contexts in which *troncamento* is obligatory, and in which it is optional or impossible.

3. Prosodic conditioning of *troncamento*

3.1. Basic observations

A closer look at the data, drawn from the C-Oral-Rom corpus and collected in interviews with two Tuscan informants, shows that *troncamento* applies to certain words obligatorily in four syntactic contexts. First, it applies without exception to the final vowel of the infinitive of any verb when followed by an enclitic pronoun (as noted in most descriptive grammars, cf., e.g., Serianni 1988: 31, 258), see (11). Second, it applies obligatorily to the final vowel of the infinitive of the modal verbs *volere* 'want', *dovere* 'have to', *potere* 'be able to', *sapere* 'be able to' when these are followed by another infinitive, as in (12) (Vogel et al. 1983: 208, 223; Cardinaletti and Shlonsky 2004: 529; this fact[4] is seldom noted in descriptive grammars, with the exception of Battaglia and Pernicone 1962: 34). In the corpus, word-final /e/ is deleted in 79 of 79 occurrences in constructions with the modal verbs *volere* 'want', *dovere* 'have', *potere* 'be able to', *sapere* 'be able to', and it is deleted in 55 of 61 occurrences with the causative verb *fare* 'make', when these are followed by an infinitive. Third, *troncamento* applies with very high frequency to the final vowel of certain nouns when they are immediately followed by another noun (Serianni 1988: 31), see (13), as well as in some, but not all, noun-noun compounds, like *quartiermastro* ‘quartermaster’ (< *quartiere mastro*). In the corpus, word-final /e/ of a noun followed by another noun is deleted in 19 of 19 occurrences of *dottore* ‘doctor’, in 13 of 14 occurrences of *professore* ‘professor’, in 6 of 6 occurrences of *signore* ‘mister’, but in 0 of 7 occurrences of *generale* ‘general’. Finally, its application is obligatory or highly frequent in the case of the adverbs *bene* ‘well’ and *male* ‘badly’ when followed by a verb or adjective they modify (cf. also Serianni 1988: 31), see (14). In the corpus, word-final /e/ is deleted in 50 of 51 occurrences of the adverb *bene*, and in 1 of 1 occurrence of the adverb *male*.

(11)

fare > far

non	*me*	*lo*	*permette*	*di*	*far.lo oggi*
not	me	it	allows	to	do.it today

‘It does not allow me to do it today.’

(12)

volere > voler

ma	*sembrano*	*non*	*voler*	*capire*
but	they.seem	not	want	understand

‘But they seem not to want to understand.’

(13)

professore > professor

saluto	*ancora*	*il*	*professor*	*Bettini*
I. greet	again	the	professor	Bettini

‘I greet professor Bettini once again.’

(14)

bene > ben

eventi	*ben*	*più*	*gravi*
events	well	more	serious

‘much more serious events.’

In contrast, *troncamento* is blocked between a subject DP and a VP, see (15), between a DP and a PP (which is not a complement of the DP) or a PP and a DP, see (16), as well as at the end of an utterance, see (17). The corpus yields no counterexamples to this generalization, and blocking of *troncamento* in these contexts has been confirmed in the informant interviews.

(15)

la	prima	[cosa	da	fare]$_{NP}$	[è	bagnare	la parte	...]$_{VP}$
la	prima	[cosa	da	*far]$_{NP}$	[è	bagnare	la parte	...]$_{VP}$
the	first	thing	to	do	is	wet	the (body.)part ...	

‘the first thing to do is to wet the body part ...’

(16)

sto insegnando	in una	[scuola	superiore]$_{NP}$ il	[linguaggio gestuale]$_{NP}$
sto insegnando	in una	[scuola	*superior]$_{NP}$ il	[linguaggio gestuale]$_{NP}$
I.am teaching	in a	school	higher	the language sign

‘I am teaching sign language in a high school.’

(17)

non	[ho	nessuna	[domanda	da	[fare]$_{VP}$]$_{NP}$]$_{VP}$
non	[ho	nessuna	[domanda	da	[*far]$_{VP}$]$_{NP}$]$_{VP}$
not	I.have	none	question	to	make

‘I have no question to ask.’

In some contexts, *troncamento* is optional. Between a verb and a following direct object *troncamento* often does not apply, as in the examples in (18) and (20), though it can apply in such contexts, as in (19) and (21).

(18)

è	difficile	[fare	delle	[distinzioni]$_{NP}$]$_{VP}$
it.is	difficult	to.make	of.the	distinctions

‘It is difficult to make distinctions.’

(19)

è	riuscito	a	[far	dei	[casini]$_{NP}$	anche	a	Montecarlo]$_{VP}$
he.is	managed	to	make	of.the	houses	also	in	Montecarlo

‘He has managed to build houses also in Montecarlo.’

(20)

perché	i	pesci	[mangiano	altri	[pesci]$_{NP}$]$_{VP}$
because	the	fishes	eat	other	fishes

‘Because the fishes eat other fishes.’

(21)

gli	americani	[mangian	tutte	[salse]$_{\text{NP}}$]$_{\text{VP}}$
the	Americans	eat	all	sauces

‘The Americans eat all kinds of sauces.’

While the informants accepted such constructions both with and without *troncamento*, the corpus data show a more differentiated picture. In a sample containing 489 instances of infinitives followed by a direct object DP or (VP-internal) PP drawn from the C-Oral-Rom corpus (taking into account only infinitives that occur at least once in the corpus in truncated form), only 35 (7 percent) showed *troncamento*. Splitting up the sample according to whether the DP or PP following the infinitive contains one or more than one lexical word (as opposed to a function word – see section 3.2 for a motivation for making this distinction), we find the final /e/ of the infinitive to be truncated in 11 percent of the cases where DP or PP contains only one lexical word. No instances of *troncamento* were found on infinitives followed by a DP or PP containing more than one lexical word. *Troncamento* is thus relatively infrequent, but licit, when the target word is followed by a short syntactic constituent, but it never applies when the target word is followed by a longer syntactic constituent.

Finally, *troncamento* applies optionally to very few nouns (number of types attested in the corpus is 5), such as *bicchiere* ‘glass’, when followed by a modifier, and to very few adjectives (8 types), such as *maggiore* ‘greater’, when followed by a word that they modify, as illustrated in (22) and (23).

(22)

il caffè	io	lo	bevo	sempre	in	[bicchier	di	[vetro]$_{\text{NP}}$]$_{\text{NP}}$
the coffee	I	it	drink	always	in	cup	of	glass

‘As to coffee, I always drink it from a glass.’

(23)

l’	elezione	di	un	[maggior	[numero	di	[parlamentari]$_{\text{NP}}$]$_{\text{NP}}$]$_{\text{AP}}$
the	election	of	a	greater	number	of	parliament members

‘the election of a greater number of members of the parliament’

Of the 108 occurrences of nouns that allow *troncamento* in this context, the final vowel is deleted in 40 cases, while it remains intact in 68 cases. Of the 61 occurrences of adjectives that allow vowel deletion in this context, *tronca-*

mento applies in 36 cases. Crucially, *troncamento* never applies to nouns and adjectives in contexts other than these; i.e., it does not apply to these nouns when they are not followed by a modifier but by a major syntactic boundary; likewise, it never applies to adjectives when they are not followed by an element which is modified by them.

3.2. *Troncamento* interacts with prosodic phrasing

At this point, it is becoming evident that the application of *troncamento* is partly determined by the syntactic structure of an utterance, or rather, by the prosodic reflection of syntactic structure. That the prosodic reflection of syntactic structure has an influence on many aspects of phonological rules and representations has been shown for a variety of languages in phonological research of the last decades (Selkirk 1978, 1980 [1972], 1984, Nespor and Vogel 1986). Many researchers assume that prosodic structure consists, like syntactic structure, of hierarchically structured constituents, where the prosodic constituents belong to a restricted set defined with recourse to some version of the prosodic hierarchy (Selkirk 1978 and subsequent research).

An influential model of the syntax-phonology interface assumes an edge-based mapping between morphosyntactic structure and prosodic structure (Selkirk 1986, Chen 1987). Under this approach, edges of syntactic constituents, such as maximal projections (henceforth 'XP'), are aligned with edges of prosodic constituents, such as phonological phrases (in the following 'PPh'), via constraints of the Alignment family (McCarthy and Prince 1993). Similarly, edges of morphological constituents, such as morphological words (henceforth 'MWd'), are aligned with edges of prosodic constituents, such as prosodic words ('PWd'), syllables or feet. Following work by Selkirk (1995), Truckenbrodt (1999) and others, I assume that the relation between morphological and syntactic structure on the one hand and prosodic structure on the other can be captured in a constraint-based model like Optimality Theory by the constraints given in (24) and (25). The constraints in (24) relate to the mapping of syntactic X-bar structure onto phonological phrase structure.

(24) Constraints on the mapping of syntactic structure onto prosodic structure (cf. Truckenbrodt 1999)

a. ALIGN XP, R: ALIGN (XP, R; PPH, R)
For each XP there is a PPh such that the right edge of XP coincides with the right edge of PPh.

b. ALIGN XP, L: ALIGN (XP, L; PPH, L)
For each XP there is a PPh such that the left edge of XP coincides with the left edge of PPh.

c. WRAPXP
Each XP is contained in a PPh.

Following paragraphs will illustrate how the interaction of these constraints derives phonological phrase structure, in cases where *troncamento* is or is not applied. Since the notion of prosodic word plays a crucial role for the analysis of *troncamento* as alternant selection developed in section 4, let me briefly comment on the mapping of morphological words onto prosodic words. By morphological word, I refer to any entity which is insertable into an X^0 slot in syntactic structure (Di Sciullo and Williams 1987; cf. also Dixon and Aikhenvald's 2002 notion of grammatical word). The mapping of morphological words onto prosodic words is subject to the constraints given in (25).

(25) Constraints on the mapping of morphological word structure onto prosodic word structure (cf. Selkirk 1995)

a. ALIGN MWD, R: ALIGN (MWD[5], R; PWD, R)
For each MWd there is a PWd such that the right edge of MWd coincides with the right edge of PWd.

b. ALIGN MWD, L: ALIGN (MWD, L; PWD, L)
For each MWd there is a PWd such that the left edge of MWd coincides with the left edge of PWd.

Note that both ALIGN MWD, R and ALIGN MWD, L can be violated, as indicated by the fact that processes of segment deletion like *troncamento*, but also processes of segment insertion or spreading (e.g., Italian *raddoppiamento*), can apply at boundaries of morphological words. Since the ranking of the constraints in (25) with respect to the alignment constraints in (24) and with respect to additional constraints that will be introduced in section 4 is not relevant for the purpose of the present paper, I will not come back to them.

A final note on the morphosyntax-prosody mapping is in order here. Much recent research has shown that alignment of boundaries of XPs and phonological phrase boundaries refers only to maximal projections of lexical heads (such as nouns, verbs and adjectives), not to maximal projections of functional heads (such as prepositions, determiners, pronouns etc.), in that boundaries of functional projections, like PP, DP, NegP, CP etc. are not related to prosodic boundaries (cf. Selkirk's 1984: 343 principle of the categorical invisibility of function words to the principles of syntax-phonology mapping, cf. also Truckenbrodt's 1999 Lexical Category Condition). Similarly, function words, unlike lexical words, in many cases do not correspond to prosodic words (Selkirk 1995). Rather, they can be prosodically integrated in a variety of ways. Though the precise conditions on prosodic integration in terms of procliticization or encliticization of function words in Italian remain to be clarified, for the sake of simplicity I assume that functional words may be parsed as syllables and feet which are adjoined to the phonological phrase on their right (procliticization; cf. Peperkamp 1997: 195–6). Integration of functional words into the phonological phrase on their left is ruled out by the constraint given in (26).

(26) ALIGN (PPH, R; PWD, R) (Selkirk 1995: 33)
For each PPh there is a PWd such that the right edge of PPh coincides with the right edge of PWd.

The hypothesis that lies at the basis of the analysis proposed here is that Italian *troncamento* is a segmental alternation whose domain is restricted to the phonological phrase (cf. Nespor 1990 for a different proposal,[6] based, however, on data from Italian as spoken in Milan), i.e., it does not apply when immediately followed by a phonological phrase boundary (cf. Meinschaefer 2006). The assumption that *troncamento* is blocked at a boundary of a phonological phrase, but applies everywhere else, can account for the observation that application of *troncamento* is obligatory in contexts like (11) to (14) above, where no right boundary of a syntactic XP interferes between the target of *troncamento* and the following word, given that the phonological phrase structure of these contexts looks like (27). Likewise, this assumption can explain why *troncamento* is blocked in contexts like (15) to (17) above, in all of which the target of *troncamento* is separated from the following word by a right XP boundary, if the phrasing for these constructions is as given in (28).

(27)

a. Syntax[7] non me lo [permette di [far**e**.lo oggi]$_{VP}$]$_{VP}$
Phonology (non me lo permette)$_{PPh}$ (di far=lo oggi)$_{PPh}$
‘It does not allow me to do it today.’

b. Syntax [sembrano non [voler**e** capire]$_{VP}$[8]]$_{VP}$
Phonology (sembrano)$_{PPh}$ (non voler capire)$_{PPh}$
‘They seem not to want to understand.’

c. Syntax [saluto ancora il [professor**e** Bettini]$_{NP}$]$_{VP}$
Phonology (saluto)$_{PPh}$ (ancora il professor Bettini)$_{PPh}$
‘I greet professor Bettini once again.’

d. Syntax [eventi [ben**e** più gravi]$_{AP}$]$_{NP}$
Phonology (eventi)$_{PPh}$ (ben più gravi)$_{PPh}$
‘much more serious events’

(28)

a. Syntax la prima [cosa da [far**e**]$_{VP}$]$_{NP}$ [è [bagnare . . .]$_{VP}$
b. Phonology la prima cosa (da far**e**)$_{PPh}$ (è bagnare . . .)$_{PPh}$
‘the first thing to do is to wet . . . ’

For the examples given in (27) and (28), a phonological phrase structure which is consistent with the assumption that *troncamento* applies within the phonological phrase, but not across phonological phrase boundaries, can be derived from the constraints given in (24) above as follows. First, it is clear that Align XP, R must be ranked high in order to rule out the possibility of having *troncamento*, and thus absence of a phonological phrase boundary, between subject and verb as in the example in (28). Note that in the tableau in (29) below, I leave out phonological phrase boundaries whose presence cannot, for the moment, be established on the basis of application of *troncamento*. For reasons of space I leave out prosodic word structure. Likewise, only candidates are considered that do not violate Align (PPh, R; PWd, R), i.e., where function words like prepositions, determiners, pronouns or negation particles are adjoined to the prosodic phrase on their right, and not on their left. Since *troncamento* was found never to apply in contexts where the target word is followed by a right boundary of an XP, the constraint Align XP, R must be undominated.

(29)

	la prima [cosa da [fare]$_{\text{VP}}$]$_{\text{NP}}$ [è bagnare . . .]$_{\text{VP}}$	Align XP, R
a. ☞	la prima cosa da fare)$_{\text{PPh}}$ (è bagnare . . .)$_{\text{PPh}}$	√
b.	la prima cosa da fare è bagnare . . .)$_{\text{PPh}}$	*

It is easy to see that in none of the constructions given in (27) above would a phonological phrase boundary be forced by Align XP, R before a word that undergoes *troncamento*. But also the constraint WrapXP must play some role in Italian, since in (27a) to (27d) *troncamento* applies obligatorily or at least with high frequency, indicating that prosodic boundaries may not be freely inserted within XPs. As can be seen in (30), WrapXP prevents insertion of additional phonological phrase boundaries within an XP, which would block *troncamento*.

(30)

	[sembrano non [volere capire]$_{\text{VP}}$]$_{\text{VP}}$	Align XP, R	WrapXP
a. ☞	sembrano non voler capire)$_{\text{PPh}}$	√	√
b.	sembrano non volere)$_{\text{PPh}}$ (capire)$_{\text{PPh}}$	√	*!

The ranking of Align XP, R and WrapXP with regard to each other cannot be established at the moment. Before we come to this point, let us first note that the third constraint, Align XP, L, likewise has a role to play in Italian. As was evidenced by the corpus data cited in section 3, *troncamento* applies optionally to the final /e/ of an infinitive when followed by a direct object DP or by a complement PP, but the application rate is low, around 10 percent. Hence, a constraint must be invoked that blocks application in the remaining 90 percent. It is easy to see that this must be Align XP, L. As shown in (31), each of the two possible phrasings, as indicated by application (in 10 percent of the cases) or non-application (in 90 percent of the cases) of *troncamento*, violates one of the two constraints.

(31)

	perché i [pesci]$_{\text{NP}}$ [mangiano altri [pesci]$_{\text{NP}}$]$_{\text{VP}}$	Align XP, R	Align XP, L	WrapXP
a.	(perché i pesci)$_{\text{PPh}}$ (mangiano)$_{\text{PPh}}$ (altri pesci)$_{\text{PPh}}$	√	√	*
b.	(perché i pesci)$_{\text{PPh}}$ (mangian altri pesci)$_{\text{PPh}}$	√	*	√

Since both ALIGN XP, L and WRAPXP are violable, they must be dominated by ALIGN XP, R, which is never violated. No categorical ranking can be established, however, for ALIGN XP, L and WrapXP. I will not discuss here the question how this kind of variation should best be represented, and whether, in addition to the length factor mentioned in section 3, syntactic factors like presence or absence of definite or demonstrative determiners, pragmatic factors like focus, or performance-related factors like speech rate, come into play here. Suffice it to say that phonological phrasing in Italian as diagnosed by application or non-application of *troncamento* is subject to variation that cannot be reduced to exhaustive ranking of categorical constraints.

4. Lexical exceptionality

4.1. Allowing *troncamento* is a property of word forms

As already shown in section 2, *troncamento* does not apply to all /e/s and to all /o/s in the same way. Only some lexemes allow *troncamento*, and they allow it only for some forms in the paradigm, but not for others. Allowing troncamento or not allowing it is thus not a property of a particular stem or of a particular suffix. Rather, it is a property of a word form, that is, of a sequence of stem and suffix. It should also be noted that the vowels deleted by *troncamento* sometimes correspond to entire suffixes (class markers in nouns, person-number markers in verbs), that sometimes they correspond to parts of suffixes, and that sometimes they are part of the stem (i.e., in adverbs).

The examples given in (32) and (33) serve to illustrate this point. The first observation to make is that while *troncamento* is allowed for the present tense forms of some verbs, like 3 sg *viene* and *vuole*, (32a), (32c), it cannot apply to 3 sg forms of other verbs that end in sonorant + /e/, like 3 sg *sale*, (32e). (In (32b), (32d) and (32e), the asterisk indicates that these forms, in addition to being unattested in the corpus, were not accepted by the informants.) The second observation is that the present tense forms of *venire* and *volere* in (32a) and (32c) allow truncation, but the past tense forms in (32b) and (32d) do not allow it. Now, this could be due to the presence of a geminate consonant (as opposed to a singleton). Still, *troncamento* (and degemination) does apply in similar cases when the mid-vowel is preceded by a geminate consonant for *hanno* > *han* and *saranno* > *saran*, see (33a) and (33b). Note, however, that *venne* > **ven* and *vanno* > *van* differ not only with regard to person, number and tense, but also with regard to quality of the deleted vowel.

(32)

a.	*vien-e*	3 sg pres	*vien*	(28 occurences)	*venire* ‘come’
b.	*venn-e*	3 sg perf	**ven*	unattested[9]	*venire* ‘come’
c.	*vuol-e*	3 sg pres	*vuol*	(112 occurences)	*volere* ‘want’
d.	*voll-e*	3 sg perf	**vol*	unattested	*volere* ‘want’
e.	*sal-e*	3 sg pres	**sal*	unattested	*salire* ‘go out’

(33)

a.	*hanno*	3 pl pres	*han*	(14 occurences)	*avere* ‘have’
b.	*saranno*	3 pl fut	*saran*	(2 occurences)	*essere* ‘be’
c.	*vanno*	3 pl pres	*van*	unattested	*andare* ‘go’

Unlike its behaviour with finite verb forms, *troncamento* applies fully productively to infinitives in at least one context, i.e., when the infinitive is followed by an enclitic pronoun. Furthermore, in other contexts infinitives may undergo *troncamento* quite productively, indicating that in the case of the infinitive it is the suffix *-re*, and not the stem or the entire word form, that determines the availability of the *troncamento* alternation. Productivity of *troncamento* on 3 pl forms ending in *-ano* and *-ono* lies somewhere between that of 3 sg, 1 sg and 1 pl forms on one hand, for which application is restricted to specific lexemes, and infinitive forms on the other: Many examples of *troncamento* on 3 pl forms are attested in the corpus, and *troncamento* is accepted by the informants for many 3 pl forms. Still, informants do not accept *troncamento* for infrequent verb forms belonging to a formal register, such as *censurano* ‘they censure’ or *certificano* ‘they certify’ (cf. also Serianni 1988: 32). The same observation can be made with regard to 1 pl forms like *siamo* > *siam* ‘we are’. While informants easily accept final vowel deletion on the 1 pl form of highly frequent verbs, like *siamo* ‘we are’, *abbiamo* ‘we have’ or *andiamo* ‘we go’, they do not accept deletion on 1 pl forms of much less frequent verbs like *censuriamo* ‘we censure’ or *certifichiamo* ‘we certify’.

Hence, if application or non-application of *troncamento* were to be marked lexically, what sort of stored forms should be marked? Because it is impossible to predict which lexemes or which suffixes by themselves license *troncamento*, it seems to be necessary to postulate that neither lexeme nor suffix is marked as [+*troncamento*], but rather that the full inflected forms are marked individually. The full forms can, however, be marked individually only if they are stored as full forms, i.e., if they are represented holistically in the lexicon, rather than being accessible only via composition from or segmentation into their component morphemes.

Now, given that the forms in question are to a large degree regular, does it really seem plausible that they are holistically represented as full forms? It has often been shown that highly frequent word forms, even when regular, can be stored in the mental lexical as full forms (e.g., Alegre and Gordon 1999, cf. Hay and Baayen 2005 for a general discussion). Less frequent word forms, on the contrary, are not stored as full forms, but are only accessible via segmentation into their component morphemes. Now, since there seems to be a relationship between the frequency of a word form and whether it undergoes the *troncamento* alternation or not, I take it that holistic representation is not only a logical prerequisite to marking full forms consisting of stem and suffix as undergoing the *troncamento* alternation, but is also plausible under cognitive aspects.

4.2. *Troncamento* as selection from multiple prosodic words in the input

The previous section has established that the entities which have to be marked as exceptionally undergoing *troncamento* are in most cases inflected forms of lexemes, and that these may be thought of as holistically stored word forms. At the same time, section 3 has shown that *troncamento* is prosodically conditioned, in that it is bound to the domain of the phonological phrase. What is more, it allows some degree of optionality, where application or non-application of *troncamento* depends, among other factors, on the length of the following constituent, and, as noted by Berruto (1987: 33), it seems to be sensitive to speech rate. The *troncamento* alternation thus appears as a prosodic alternation, i.e., as an alternation between two forms of a prosodic word. Unlike many other prosodic alternations, however, *troncamento* is not productive; rather, certain holistically stored word forms have to be marked for this alternation (in this respect, it is similar to French *liaison*, cf. Post 2000).

I think that in a modular, constraint-based system like OT this state of affairs is best represented as an alternation between listed forms whose surface distribution is accounted for by constraints on the output, i.e., as a kind of relation that is often termed ‘allomorphy’ (cf. Mascaró 1996, Kager 1996, Perlmutter 1998). However, the listed forms from which the grammar selects the optimal form must be prosodic words, not morphological words. The alternation is prosodic (or postlexical in the sense of Lexical Phonology, cf. Kiparsky 1982), in that it refers to the prosodic reflection of syntactic structure, into which morphological words cannot be directly inserted; it is only the prosodic projections of morphological words, i.e., prosodic words, that can be inserted into prosodic structure.

For a form like *viene* in (32) above, I therefore assume an underlying representation as given in (34), with one holistically stored morphological word in the input, listed together with two alternative prosodic word representations, one truncated prosodic word form and one non-truncated prosodic word form, which are likewise stored in the lexicon.

(34) *viene* 3 sg pres 'he/she/it comes'

Lexicon

Phonology	{	(v i e n e)$_{\text{PWd}}$	
		(v i e n)$_{\text{PWd}}$	}
Syntax		[[v i e n][e]]$_{\text{MWd}}$	

Note that (34) presupposes that both the morphological word [*viene*]$_{\text{MWd}}$ and the corresponding prosodic words (*viene*)$_{\text{PWd}}$ and (*vien*)$_{\text{PWd}}$ are stored in the lexicon, rather than the form *viene* being constructed from a stored stem *vien-* and stored suffix *-e* by the grammar, and then being mapped onto a prosodic word via general constraints on the syntax-prosody interface. The analysis of *troncamento* as selection from multiple prosodic words, if correct, implies that what is stored in the lexicon is not only morphemes with their segmental phonological and their morphological specifications, but also, where not predictable from general principles, prosodic structure, at least up to the level of the prosodic word. It is an open question whether prosodic structure above the word could also be represented in the input. The observation, however, that *troncamento* on nouns and adjectives is often found only in certain collocations, such as *maggior parte* < *maggiore parte* 'greater part' or *sal marino* < *sale marino* 'sea salt' lets it appear possible that this may be the case. Certain observations about French liaison seem to point into a similar direction (Bybee 2001). In non-parallel phonological models such as Lexical Phonology (Kiparsky 1982), it has long been noted that postlexical alternations fall into at least two classes: on the one hand categorical, exceptionless rules, such as Spanish /s/-aspiration, and on the other hand optional and lexically restricted rules, such as French liaison (Kaisse 1985, Hayes 1990) or Italian *troncamento*. An alternative to the analysis proposed here, where mappings from morphological to prosodic words are evaluated in a parallel, monostratal OT-grammar, might thus carry over the modular architecture of Lexical Phonology into a constraint-based OT-model, in the way proposed in Stratal

OT (Bermúdez-Otero 2007). For reasons of space, I shall not discuss this point further.

4.3. Constraints on alternant selection

If the lexical representation of a form comprises multiple inputs, the surface distribution of input alternants is expected to be governed by markedness constraints (McCarthy and Prince 1994). For the case at hand, the question that needs to be addressed is therefore which markedness constraints select the truncated form within phonological phrases, and why the full form surfaces at phonological phrase boundaries. In other words, what is the functional motivation for final vowel deletion? And what is the motivation for not deleting the vowel at a boundary of a phonological phrase? In fact, the interesting aspect of representing *troncamento* as alternant selection is precisely this: that it allows us to gain a deeper understanding of the driving forces behind this alternation.

To account for the absence of the truncated form at phonological phrase boundaries, I propose an alignment constraint which is a variant of (26) above, which was said to force procliticization, rather than encliticization, of function words. The constraint that I take to force selection of the nontruncated alternant is given in (35). This constraint requires the right edge of a phonological phrase to coincide with the right edge of a morphological word.

(35) ALIGN (PPH, R; MWD, R)
For each PPh there is a MWd such that the right edge of PPh coincides with the right edge of MWd.

The constraint in (35) is violated when right edge of the phonological phrase and right edge of the morphological word do not coincide, i.e., when the morphological word is longer or shorter than its prosodic equivalent. This constraint blocks segment deletion processes like *troncamento*, segment insertion, and segment spreading processes like *raddoppiamento* that occur at word edges, from applying at phonological phrase boundaries.

In fact, such a constraint would categorically block the occurrence of all deletion and insertion processes at edges of phonological phrases, a result which is incompatible with the observation that Italian *raddoppiamento*, i.e., consonant gemination at the left edge of a word, does occur in contexts where a phonological phrase boundary must be assumed to interfere (cf. Agostiniani 1992, D'Imperio and Gili Fivela 2003). A more differentiated picture emerges, however, if one assumes that (35) is dominated by higher ranked constraints, such as FT-BIN, requiring that feet are binary. Now, FT-BIN is also the driving force for *raddoppiamento*, whose effect is to make a stressed word-final syllable closed, i.e., heavy, thereby enabling it to be parsed as a binary foot. Crucially, the constraint in (35) would thus not be able to block *raddoppiamento* at the edge of a phonological phrase precisely because *raddoppiamento* is forced by a constraint ranked higher than (35).

Coming back to *troncamento*, in positions internal to a phonological phrase the truncated form emerges as the optimal, i.e., unmarked candidate, at least when such a form is contained in the set of inputs. Whichever constraint makes the truncated alternant 'better' than the full form must therefore be dominated by (35); otherwise, the full form would never surface. Crosslinguistically, processes of final vowel reduction and deletion are often motivated rhythmically (Kager 1997). In particular, according to Kager (1997) a driving force involved in syncope may be what he terms "exhaustivity of metrical parsing", i.e., a pressure towards avoiding metrically unparsed or unfooted syllables. To understand whether such a rhythmic motivation may also motivate selection of the truncated alternant in Italian *troncamento*, let me first present some basic observations about footing and stress assignment in Italian.

With regard to Italian stress I take as relevant (among others) the constraints given in (38) to (40) below, and I make the following general assumptions (cf., e.g., D'Imperio and Rosenthall 1999, Jacobs 1994, Meinschaefer 2006, Vincent 1988). Feet are moraic trochees. Main stress is on the rightmost foot (ALIGN [FOOT, R; PWD, R]). I will not consider secondary stress and rhythm; I rather assume that Italian has maximally one foot per word (PARSE-σ, see (42) below, is dominated by ALIGN (FOOT, R; PWD, R)). The final syllable of a word is extrametrical (NON-FINALITY dominating ALIGN (FOOT, R; PWD, R)). The unmarked pattern is thus penultimate stress in words with a heavy penultima and antepenultimate stress otherwise, see (36).

(36) Unmarked pattern
  a. Heavy penultima: paroxytonic stress
    espérto 'expert', *leopárdo* 'leopard', *enórme* 'enormous'
  b. Light penultima: proparoxytonic stress
    líbero 'free', *último* 'last', *védono* 'they see'

Many words are, however, lexically marked for stress on the penultimate, see (37). In particular, theme vowels of verb forms are lexically marked for stress. Lexical stress is always preserved (IDENT STRESS I-O).

(37) Lexical stress

a. on the penultima: paroxytonic stress with light penultima (the vowel is lengthened by rule)
domá[:]*ni* 'tomorrow', *leggé*[:]*ro* 'light', *tanní*[:]*no* 'tannin'

b. on the ultima: oxytonic stress (no lengthening applies)
cittá 'city', *virtú* 'virtue', *caffé* 'coffee'

(38) Constraints on foot form

a. FT-BIN
Feet are binary (under moraic analysis).

(39) Constraints on foot position

a. ALIGN (FOOT, R; PWD, R)
Stress is rightmost in the prosodic word.

b. NON-FINALITY
A foot may not be final in the morphological[10] word.

(40) Correspondence constraints

a. IDENT STRESS I-O
Lexical stress is preserved in the output.

b. IDENT STRESS 3SG-3PL
In a 3 pl form of a verb stress is located in the same position as in the 3 sg form of that verb. [11]

Note that ALIGN (FOOT, R; PWD, R) must be dominated by FT-BIN and NON-FINALITY; otherwise, stress would be on the ultima or penultima, but not on the antepenultima. ALIGN (FOOT, R; PWD, R) must also be dominated by IDENT LENGTH I-O (see below); otherwise, stress would surface on the (lengthened) penultima rather than on the antepenultima. Finally, IDENT STRESS I-O must dominate FT-BIN, NON-FINALITY and IDENT LENGTH I-O; otherwise, lexically prespecified stress could never surface on the ultima or on a lengthened penultima.

A comparison of the full forms and of the truncated forms in (41) with regard to foot structure now shows the following differences. On one hand, full forms have stress on the penultimate or antepenultimate syllable, so that syllables to the right of the main foot remain unparsed. In the paroxytonic forms in (41a), (41c) and (41d), the stressed vowel is lengthened. The truncated forms, on the other hand, have final or penultimate stress. In the truncated forms there is one syllable less which remains unfooted, i.e., metrically unparsed, than in the full form, and no vowel lengthening applies.

(41)

	underlying form	full form	truncated form	
a.	/mandʒano/	(mán).dʒa.no	(mán).dʒan	'they eat',V
b.	/minóre/	mi.(nó:).re	mi.(nór)	'minor', Adj
c.	/bene/	(bé:).ne	(bén)	'well', Adv

As can be inferred from (41), final vowel deletion has thus two effects: First, it reduces the number of unparsed syllables in the examples given in (41). Second, it obviates vowel lengthening in (41b) and (41c), because as a result of final vowel deletion the stressed penultima is closed, and thus heavy. I therefore regard the constraints given in (42) as relevant to selecting the truncated alternant as the optimal candidate.

(42)

a.	IDENT LENGTH I-O	No lengthening!
b.	PARSE-σ	All syllables are parsed into feet.[12]

Since *troncamento* does not apply productively, but only to those word forms which dispose of multiple inputs, including one truncated and one non-truncated alternant, the two constraints motivating deletion must be dominated by a constraint against productive vowel deletion given in (43). If MAX-VOWEL were not ranked high in the hierarchy, then the Gen component would freely generate output candidates with deleted vowels even for forms whose input does not contain a truncated alternant. These output candidates would then be selected as optimal by IDENT LENGTH I-O and PARSE-σ, contrary to the attested restricted productivity of *troncamento*.

(43) MAX-VOWEL No vowel deletion!

4.4. Some analyses

Now let me briefly illustrate how the constraints given above lead to selection of the optimal form, truncated or non-truncated. For the sake of clarity I leave out undominated constraints wherever possible, and I begin with the simplest case, the two-syllable adverb *bene*; see tableau (45). The input word form has two alternants. The morphological word occurs in a syntactic context where it is not immediately followed by a syntactic boundary, as in (44). As shown in (45), the non-truncated form violates both Parse-σ and Ident Length I-O, and also Align (F, R; PWd, R). Therefore, the truncated alternant is selected by Eval.

(44)

Syntax [eventi [bene più gravi]$_{\text{AP}}$]$_{\text{NP}}$
Phonology (eventi)$_{\text{PPh}}$ (ben più gravi)$_{\text{PPh}}$
‘much more serious events’

(45)

{(bene)$_{\text{PWd}}$, (ben)$_{\text{PWd}}$} [bene]$_{\text{MWd}}$ ...]$_{\text{XP}}$	Ft-Bin	Non-Final	Parse-σ	Ident Lenght I-O	Align (F, R; PWd, R)
a. (bé:).ne	√	√	!*	*	*
b. ☞ (bén)	√	√	√	√	√

Now let us consider an expression where *bene* is followed by an XP boundary in syntactic structure. (47) starts with the same set of multiple inputs. But in this case the morphological word is followed by an XP boundary in syntactic structure, corresponding to a phonological phrase boundary in prosodic structure, as in (46). Therefore, the truncated alternant is ruled out by the undominated constraint Align (MWd, R; PPh, R).

(46)

Syntax [lo sanno bene]$_{\text{VP}}$ gli [insegnanti]$_{\text{NP}}$
Phonology (lo sanno bene)$_{\text{PPh}}$ (gli insegnanti)$_{\text{PPh}}$
‘the teachers know it well’

(47)

{(bene)$_{\text{PWd}}$, (ben)$_{\text{PWd}}$} [bene]$_{\text{MWd}}$]$_{\text{XP}}$	Align (PPh, R; MWd, R)	Non-Final	Parse-σ	Ident Lenght I-O	Align (F, R; PWd, R)
a. ☞ (bé:).ne	√	√	*	*	*
b. (bén)	*	√	√	√	√

Now consider an input like *meno* ‘less’ that, unlike *bene* ‘well’ and *male* ‘badly’, does not dispose of two alternants, as in (48). In this case, the truncated alternant is not contained in the input; therefore, the truncated form in (48b) is ruled out by MAX-V.

(48)

[[meno]$_{\text{MWd}}$ . . .]$_{\text{XP}}$	MAX-V	FT-BIN	NON-FINAL	PARSE-σ	IDENT LENGHT I-O	ALIGN (F, R; PWD, R)
a. ☞ (me:).no	√	√	√	*	*	*
b. (men)	!*	√	√	√	√	√

Also for an input form with lexical stress on the penultima, such as the adjective *minóre* ‘minor’, the truncated form is selected as optimal candidate, while the non-truncated form is ruled out by PARSE-σ, IDENT LENGTH I-O and also by ALIGN (F, R; PWD, R), see (49).

(49)

{(minore)$_{\text{PWd}}$, (minor)$_{\text{PWd}}$} [minór-e]$_{\text{MWd}}$...]$_{\text{XP}}$	IDENT STRESS I-O	FT-BIN	NON-FINAL[13]	PARSE-σ	IDENT LENGTH I-O	ALIGN (F, R; PWD, R)
a. mi.(nó:).re	√	√	√	*!*	*	*
b. ☞ mi.(nór)	√	√	√	*	√	√

Finally, (50) shows that the truncated alternant is likewise selected for 3 pl verb forms like *mangiano*, unless it is immediately followed by a phonological phrase boundary.

(50)

{(mandʒano)$_{\text{PWd}}$, (mandʒan)$_{\text{PWd}}$} [mandʒ-ano]$_{\text{MWd}}$...]$_{\text{XP}}$	FT-BIN	IDENT STRESS 3SG-3PL	NON-FINAL	PARSE-σ	ALIGN (F, R; PWD, R)
a. (mán)dʒa.no	√	√	√	*!*	**
b. ☞ (mán).dʒan	√	√	√	*	*
c. (man).(dʒán)	√	!*	√	√	√

Note that here one might expect a rightward stress shift in the truncated form relative to the full form, due to ALIGN (F, R; PWD, R), as in (50c). In fact, since NON-FINALITY is defined with regard to the morphological word and not the prosodic word (see above and note 10), NON-FINALITY cannot be the relevant constraint in ruling out the (incorrect) candidate in (50c). This candidate is ruled out, however, by a higher ranked constraint, IDENT STRESS 3SG-3PL, which is needed on independent grounds, see note 11.

5. Conclusion

This study has given a detailed account of the patterns of lexical exceptionality that occur for final vowel deletion, or *troncamento*, in Standard Italian as spoken in Tuscany. While *troncamento* presents some phonological regularities, in that it applies regularly within phonological phrases, but not when immediately followed by a phonological phrase boundary, it does not apply to all lexemes with the same productivity. Crucially, it is not lexemes or suffixes that are marked as undergoing *troncamento*, but rather fully inflected word forms.

These patterns of lexical exceptionality can be accounted for under the assumption that the forms undergoing *troncamento* are lexically represented as having two alternants, one truncated and one non-truncated alternant. Since *troncamento* presents some properties of a prosodic alternation, I have proposed that the two stored alternants are prosodic words (rather than morphological words). Selection of one of the two alternants has been assumed to be governed by hierarchically ranked constraints in an Optimality Theoretic grammar. In the analysis proposed here, selection of the truncated alternant is motivated by a constraint against unfooted syllables as well as a constraint against vowel lengthening (where the non-truncated alternant is paroxytonic), while selection of the truncated alternant at phonological phrase boundaries is blocked by a constraints requiring that the edge of the phonological phrase must coincide with the edge of the morphological word.

Acknowledgements

Previous versions of this work have been presented at the conference on *Variation and change in Phonology and Phonetics*. Potsdam 2004, at the *Second Old World Conference in Phonology*, Tromsø 2005 and at the linguistics colloquium of the *UMR 7023 Structures formelles du language*, University of Paris 8. I would like to thank the audiences as well as Aditi Lahiri and Christoph Schwarze for discussion of various aspects of this study. I would also like to thank two anonymous reviewers for helpful comments on an earlier version of this manuscript. Any remaining errors and omissions are my own. This work has been supported by the German Science Foundation through its *Sonderforschungsbereich 471*, University of Konstanz.

Notes

1. The C-Oral-Rom corpus (Cresti and Moneglia, eds., 2005) is a corpus of spontaneously spoken Italian, compiled by the LABLITA at the University of Florence. It contains 150 hours of recording of spontaneous speech, corresponding to roughly 300,000 words. Both formal and informal speech are covered, and the corpus contains monologues as well as conversations of two or more speakers. The material is accessible in orthographic transcription, aligned with the original recordings.
2. Here and in the following examples, prosodic annotation as found in the C-Oral-Rom corpus is omitted, since it is not clear how the prosodic annotation in the corpus relates to the categories of prosodic phonology that are employed in the analysis presented here (along the lines of Selkirk 1978, Nespor and Vogel 1986 and subsequent work). It can be said, however, that "non-conclusive breaks", indicated in the corpus by means of the symbol '/' (Moneglia 2005), do not stand in a systematic relation to phonological phrases as they are conceived of here.
3. In an analysis of Spanish word structure, Harris (1996) proposes that all lexemes, including adverbs, may be taken to end in a form class marker. Though Spanish and Italian may be said to share many structural properties, it is not clear at the moment whether Harris' analysis of Spanish may be carried over to Italian.
4. This is also confirmed by our own corpus data. For the 79 occurrences of the modal verbs *dovere* 'have to', *volere* 'want', *potere* 'be able to', *sapere* 'be able to/know', truncation of the final vowel of the infinitive of the modal verb is attested in all cases.
5. The constraints (25a) and (25b) are identical to Selkirk's (1995) Word Alignment Constraints. While Selkirk's original formulation relates to lexical words ('Lex') as opposed to function words ('Func') I have modified it so as to refer to morphological words ('MWd') in general, where the constraint is precluded from reference to function words by the more general principle of the invisibility of functional structure to the principles of morphosyntax-phonology mapping.
6. Nespor (1990) assumes that in Standard Italian as spoken in Milanese, the prosodic domain of *troncamento* is the clitic group. In addition, she assumes that *troncamento* applies in an optional and unconstrained manner in the intonational phase.
7. Since functional structure plays no role in the morphosyntax-prosody mapping (Selkirk 1984, Truckenbrodt 1999), in the syntactic representations given

throughout this text, only XPs projected from lexical heads are indicated, i.e., VPs, NPs and APs.

8. For sequences of modal or causative verb + infinitive, I assume a syntactic representation where the two V heads combine to form a V° (Rizzi 1978; cf. Cardinaletti and Shlonsky 2004 for discussion and for an alternative proposal that is likewise compatible with the phonological phrase structure assumed here). Note that the special syntactic structure of these V+V sequences is the reason why, in expressions like *voler fare* 'want to do', *troncamento* is obligatory, while it is optional in V+DP sequences like *volere un gelato* 'want ice cream'.
9. The forms for which application of *troncamento* would be conceivable, but which do not occur in the corpus in truncated form, have been presented to two native informants. The 3 sg forms and the (lexicalized) 3 pl forms with geminate consonant for which truncation is unattested in the corpus were not accepted by the informants, with the exception of the form *fanno* > *fan* 'they do' and *vanno* > *van* 'they go'. As to the (productively formed) 3 pl forms ending in *-ano* and *-ono*, in contrast, the informants accepted many forms which are not attested in the corpus.
10. Adjectives like *minóre* that surface as a truncated alternant *mi(nór)*, as well as truncated infinitives in general, pose a problem for the assumption that a foot may not be final. At the level of the prosodic word, the foot is in fact final in a form like *mi(nór)*; at the level of the morphological word, however, the foot is not final, given that the morphological word is /minór-e/. At the same time, for independent reasons NON-FINALITY needs to be ranked relatively high in Italian (dominated by IDENT STESS I-O, however, to account for words with oxytonic stress like *vir(tú)* 'virtue'). To account for this seemingly odd situation, I propose to define NON-FINALITY with regard to the morphological word rather than the prosodic word. In this way, NON-FINALITY leads to the desired effect, i.e., extrametricality, for all forms for which the morphological word is identical to the prosodic word, i.e., for all non-truncated forms, but it does not ban *troncamento* from applying to words with (lexical) stress on the penultima. NON-FINALITY is thus a constraint which relates to the mapping of morphological structure and rhythmic structure rather than to rhythmic structure by itself. It could be formulated as a constraint stating that (certain) inflectional endings are extrametrical.
11. This constraint is needed to account for superproparoxytonic stress in 3 pl forms of verbs like *fabbricare* 'fabricate', *imitare* 'imitate', *calcolare* 'calculate'. In the 3 sg, the forms of these verbs show regular antepenultimate stress, as in *(fáb)brica*, *(ími)ta*, *(cál)cola*. In the 3 pl, stress is on the preantepenultimate syllable, as in *(fáb)bricano*, *(ími)tano*, *(cál)colano*. This pattern cannot be derived in a rule-based or constraint-based way. Since the generalization that in 3 pl verb forms stress is on the same syllable as in the corresponding 3 sg verb forms holds across the Italian verb system, I propose that the 3sg-3pl identity con-

straint is responsible for preantepenultimate stress in these forms. Note that this constraint dominates both PARSE-σ and ALIGN (F, R; PWD, R). An alternative would be to exceptionally mark the 3 pl suffix $\{$*ano, ono, ino*$\}$ as extrametrical.

12. PARSE-σ is violated by an unparsed syllable to the right or to the left of the main foot. Still, since I take PARSE-σ to be dominated by ALIGN (F, R; PWD, R), a prosodic word has only one foot.
13. NON-FINALITY has to dominate PARSE-σ and ALIGN (F, R; PWD, R). Otherwise, vowel final words would always be stressed on the penultima, and consonant-final words, like *épsilon*, *céllofan*, etc., would be oxytonic, contrary to fact.

References

Agostiniani, Luciano
1992 Su alcuni aspetti del 'rafforzamento sintattico' in Toscana e sulla loro importanza per la qualificazione del fenomeno in generale. *Quaderni del dipartimento di linguistica* 3: 1–28.

Alegre, M. and P. Gordon
1999 Frequency effects and the representational status of regular inflections. *Journal of Memory and Language* 40: 41–61.

Battaglia, Salvatore and Vincenzo Pernicone
1962 *Grammatica italiana*. Torino: Loescher.

Bermúdez-Otero, Ricardo
2007 Morphological structure and phonological domains in Spanish denominal derivation. In *Optimality-Theoretic Studies in Spanish Phonology*. Fernando Martínez-Gil and Sonia Colina (eds.), 278–311. Amsterdam: Benjamins.

Berruto, Gaetano
1987 *Sociolinguistica dell'italiano contemporaneo*. Roma: La Nuova Italia Scientifica.

Bybee, Joan
2001 Frequency effects on French liaison. In *Frequency and the emergence of linguistic structure*. Joan Bybee and Paul Hopper (eds), 337–359. Amsterdam: Benjamins.

Cardinaletti, Anna and Ur Shlonsky
2004 Clitic Positions and Restructuring in Italian. *Linguistic Inquiry* 35: 519–557.

Chen, Matthew Y.
1987 The syntax of Xiamen tone sandhi. *Phonology Yearbook* 4: 149.

Chomsky, Noam and Morris Halle
1968 *The sound pattern of English*. New York: Harper and Row.

Cresti, Emanuela and Massimo Moneglia (eds)
2005 *C–ORAL–ROM Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam: Benjamins.

Di Sciullo, Anna–Maria and Edwin Williams
1987 *On the definition of word*. Cambridge, Mass.: MIT Press.

D'Imperio, Mariapaola and Barbara Gili Fivela
2003 How many levels of phrasing? Evidence from two varieties of Italian. *In Phonetic Interpretation. Papers in Laboratory Phonology VI*. John Local, Richard Ogden and Rosalind Temple (eds), 130–144. Cambridge: Cambridge University Press.

D'Imperio, Mariapaola and Sam Rosenthall
1999 Phonetics and phonology of main stress in Italian. *Phonology* 16: 1–28.

Dixon, R.M.W. and Alexandra Aikhenvald
2002 Word. A typological framework. In *Word. A cross–linguistic typology*. R.M.W. Dixon and Alexandra Aikhenvald (eds), 1–41. Cambridge: Cambridge University Press.

Harris, James W.
1996 The syntax and morphology of class marker suppression in Spanish. In *Grammatical theory and Romance languages*. Karen Zagona (eds), 99–122. Amsterdam: Benjamins.

Hay, Jennifer and Harald Baayen
2005 Shifting paradigms: gradient structure in morphology. *Trends in Cognitive Sciences* 9: 342.

Hayes, Bruce
1990 Precompiled Phrasal Phonology. In *The Phonology-Syntax Connection*. Sharon Inkelas and Draga Zec (eds.), 85-108- Chicago: The University of Chicago Press.

Inkelas, Sharon
1998 The theoretical status of morphologically conditioned phonology. A case study of dominance effects. *Yearbook of Morphology* 1997: 121–155.

Inkelas, Sharon, Orgun, C. Orhan and Zoll, Cheryl
1997 The implications of lexical exceptions for the nature of the grammar. In *Derivations and constraints in phonology*. Iggy Roca (ed.), 393–418. Oxford: Oxford University Press.

Itô, Junko and Armin Mester
1995 Japanese phonology. In *Handbook of phonological theory*. John A. Goldsmith (ed.), 817–838. Cambridge: Blackwell.

Jacobs, Haike
1994 Catalexis and Stress in Romance. In *Issues and Theory in Romance*

Linguistics. Michael Mazzola, L. (ed.), Washington DC: Georgetown University Press.

Kager, René
1996 On affix allomorphy and syllable counting. In *Interfaces in Phonology*. Ursula Kleinhenz (ed.), 155–171. Berlin: Akademie Verlag.

Kager, René
1997 Rhythmic vowel deletion in Optimality Theory. In *Derivations and constraints in phonology*. Iggy Roca (ed.), 463–499. Oxford: Oxford University Press.

Kaisse, Ellen
1985 *Connected Speech*. The Interaction of Syntax and Phonology. New York: Academic Press.

Kiparsky, Paul
1982 Lexical Phonology and Morphology. In *Linguistics in the Morning Calm*. In-Seok Yang (ed.), 1-91. Seoul: Hanshin.

Mascaró, Joan
1996 External allomorphy and contractions in Romance. *Probus* 8: 181–205.

McCarthy, John and Alan Prince
1993 Generalized Alignment. *Yearbook of Morphology* 1993: 79–153.

McCarthy, John and Alan Prince
1994 The Emergence of the Unmarked: Optimality in Prosodic Morphology. In: M. González (ed.) *Proceedings of Northeast Linguistics Society (NELS)*. Vol 24, 333–379. Cambridge: GLSA.

Meinschaefer, Judith
2006 The interface of morphology and phonology in the light of accentual patterns in Spanish and Italian verb forms. Paper presented at the *12th International Morphology Meeting*. Budapest, Mai 2006.

Meinschaefer, Judith
2006 *Troncamento* and the phonological phrase in Florentine Italian. *Rivista di Linguistica*, 279-308.

Moneglia, Massimo
2005 The C–Oral–Rom resource. In *C–ORAL–ROM Integrated reference corpora for Romance languages*. Emanuela Cresti and Massimo Moneglia (eds), 1–70. Amsterdam: Benjamins.

Nespor, Marina
1990 Vowel deletion in Italian. The organization of the phonological component. *The Linguistic Review* 7: 375–398.

Nespor, Marina and Irene Vogel
1986 *Prosodic phonology*. Dordrecht: Foris.

Pater, Joe
2000 Nonuniformity in English stress. The role of ranked and lexically specific constraints. *Phonology* 17: 237-274.

Peperkamp, Sharon
1997 *Prosodic words*. Ph. D. thesis, Amsterdam: Universiteit van Amsterdam, Faculteit der Letteren.

Perlmutter, David M.
1998 Interfaces. Explanation of allomorphy and the architecture of grammars. In *Morphology and its relation fo phonology and syntax*. Steven G. Lapointe, Diane K. Brentari and Patrick Farrell (eds), 307–338. Stanford: CSLI.

Post, Brechtje
2000 Pitch accents, liaison and the phonological phrase in French. *Probus* 12: 127–164.

Prince, Alan and Smolensky, Paul
2004 [1994] *Optimality Theory. Constraint interaction in generative grammar*. Oxford: Blackwell.

Rizzi, Luigi
1978 A restructuring rule in Italian syntax. In *Transformational studies in European language*. Jay Keyser (ed.), Cambridge, Mass.: MIT Press.

Rosati, Luca
2001 Variabilità della pronuncia e variabilità sociale in Italia. I fenomeni di sandhi nelle indagini recenti. *Studi Italiani di Linguistica Teorica e Applicata* 30: 455–488.

Schane, Sanford
1973 The Formalization of Exceptions in Phonology. In *The Formal Analysis of Natural Languages*. Maurice Gross, Morris Halle and M.P. Schuetzenberger (eds), 63–70. The Hague: Mouton.

Selkirk, Elisabeth
1978 On prosodic structure and its relation to syntactic structure. In *Nordic Prosody*. Thomas Fretheim (ed.), 111–140. Trondheim: TAPIR.

Selkirk, Elisabeth
1980 [1972] *The phrase phonology of English and French*. New York: Garland.

Selkirk, Elisabeth
1984 *Phonology and Syntax. The relation between sound and structure*. Cambridge, Mass.: MIT Press.

Selkirk, Elisabeth
1986 On derived domains in sentence phonology. *Phonology Yearbook* 3: 371–405.

Selkirk, Elisabeth
1995 The prosodic structure of function words. In *Papers in Optimality Theory*. Jill N. Beckman, Laura Walsh Dickey and Susanne Urbanczyk (eds), 439–469. Amherst, Mass.: GLSA.

Serianni, Luca
1988 *Grammatica italiana. Italiano commune e lingua letteraria*. Torino: UTET.

Truckenbrodt, Hubert
1999 On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30: 219–255.

Vincent, Nigel
1988 Non–Linear Phonology in Diachronic Perspective: Stress and Word–Structure in Latin and Italian. In *Certamen Phonologicum. Papers from the 1987 Cortona Phonology Meeting*. Pier Marco. Loporcaro Bertinetto, Michele (eds), Torino: Rosenberg and Sellier.

Vogel, Irene, Marina Drigo, Alessandro Moser and Irene Zannier
1983 La cancellazione di vocale in italiano. *Studi di Grammatica Italiana* 12: 189–230.

Zonneveld, Wim
1978 *A formal theory of exceptions in generative phonology*. Dordrecht: Foris.

On the distribution of dorsals in complex and simple onsets in child German, Dutch and English

Ruben van de Vijver, Barbara Höhle & Susan Ott

The aim of this paper is to argue that there is substantive variation between child language and adult language, contrary to what is often assumed in acquisition theories.[1] Our data show that some children have a distribution of dorsals–allowed in clusters, but not in singletons–that is typologically unattested. The grammar needed to describe the language of the children is never needed to describe any adult language. In other words, there is appears to be variation in the substance of child and adult grammars.

In Optimality Theory (Prince and Smolensky, 1993/2004), a grammar is a hierarchical ranking of a set of universal constraints.[2] A stage in child language is a ranking of these constraints and should correspond to an adult language (Levelt and van de Vijver, 2004). We call this kind of variation between child grammars and adults grammars unsubstantive variation. An example of such unsubstantive variation comes from the acquisition of syllable types. Children acquiring Dutch have to learn that syllables with complex onsets and complex codas are allowed. The grammar of the children goes through a stage in which only simple onsets and codas are allowed. This stage can be described with the same constraints–markedness constraints against complex syllable margins outrank faithfulness constraints–that are necessary to describe the final stage–faithfulness constraints now outrank markedness constraints. Moreover, the stage in which only simple margins are allowed corresponds to the grammar of Thargari (Levelt and van de Vijver, 2004).

A substantive difference between child language and adult language, on the other hand, is found when child language has properties not found in any adult language. No adult language requires codas and, as a consequence, there is no constraint or constraint ranking that would cause epenthesis of consonants in case a syllable ends in a vowel. If a child language would require codas and we were to set up a constraint or a ranking that would ensure epenthesis of codas in CV syllables, there would be no typological justification. With respect to segments, there are no adlt languages in which segments that are allowed in complex onsets are not alowed in simple onsets. Yet, this is what we are faced with in the data discussed in this paper. We will show that it is possible to account for the data with positional markedness (Zoll, 2004). Even though the distribution can be explained elegantly in this way, a problem remains: there are no adult languages in which dorsals are distributed in the same way.

1. Introduction

Dorsal is the last Place of Articulation (PoA) feature to be acquired in onsets (Levelt, 1994; Fikkert and Levelt, 2006) and often substituted by coronals (Bernhardt and Stemberger, 1998; Levelt, 1994; Fikkert and Levelt, 2006; Smith, 1973; Barlow, 1997; Ott, 2004; Höhle et al., 2006). The focus of this paper is on the distribution of dorsals in children acquiring German, Dutch and English. Their pattern suggest that dorsals are partially acquired, as they are allowed only in complex onsets (German), or only immediately before unstressed sonorant consonants or vowels (Dutch, English). Such a distribution is unattested typologically and therefore it raises the question to what extent constraints are universal.

Differences between child language and their adult targets are explained as differences in ranking of the *same* set of constraints (Gnanadesikan, 2004; Levelt and van de Vijver, 2004). If children have not yet acquired dorsals, but dorsals are allowed in the target language, this is seen as the result of a difference in ranking of the constraints in (1).

(1) *Markedness and Faithfulness for dorsals.*
   a. *DORSAL Segments have no dorsal PoA.
   b. IDENT(DORSAL) Segments with a dorsal PoA in the input have a dorsal PoA in the output.

In German, Dutch and English dorsals are allowed in the output, therefore their ranking is assumed to be: IDENT(DORSAL) $\gg$ *DORSAL (2-a). Children, who do not have dorsals in their output yet, have the reverse ranking (2-b).

(2) *Ranking of the Dutch target /ku/ 'cow'*

a. *Adult ranking*

/ku/	Ident(Dorsal)	*Dorsal
☞ [ku]		*
[tu]	*	

b. *Child ranking*

/ku/	*Dorsal	Ident(Dorsal)
[ku]	*	
☞ [tu]		*

This approach correctly predicts the typology of dorsals as attested in adult languages: either they are allowed everywhere (e.g. German, Dutch and English (2-a)) or nowhere (e.g. Haitian (2-b)). Unfortunately, the distribution of dorsals during acquisition is apparently more complicated than the all-or-nothing state of affairs suggested in (2-a)–(2-b).

Some German children allow dorsals in complex onsets, but not in simple onsets. They realize a word such as *Kleid* 'dress' /klaɪt/ target-like as klaɪt, but a word such as *Kuh* 'cow' /ku/ untarget-like as *tu*. Another pattern is found in children acquiring Dutch and English. They allow dorsals in complex onsets and before unstressed vowels, but not before stressed vowels. A Dutch word like *clown* [klaun] 'clown' is realized as [klaun], the dorsals a word like *kikker* [ˈkɪkər] 'frog' are realized as [t] before a stressed vowel and as [k] before an unstressed vowel: [ˈtɪkər]. dorsals are realized as dorsals in complex onsets (i.e. before unstressed sonorants): *clown* [klaun] 'clown' is realized as [klaun].

Such a distribution of dorsals is not found, as far as we know, in any language. This raises two questions, which are addressed in this paper:

(3) *Questions*

a. How can the distribution of dorsals be explained?

b. What are the typological predictions of this solution and what are its consequences for continuity?

This paper is organized as follows. The data are presented in section 2 and analyzed in 3. The consequences for continuity are explored in 4. A summary and conclusions are given in 5.

2. Data

We will discuss data of a group of children acquiring German Höhle, van de Vijver and Ott (2006), of a child acquiring English (Bills and Golston, 2002) and of one acquiring Dutch. All these children exhibit a distribution of dorsals that is not found in any adult target.

2.1. German

The German data come from a study by Ott (2004); Höhle et al. (2006). She collected data from a total of 12 children with a delayed acquisition, 6 of which consistently replaced single dorsals by single coronals.[3] All of these 6 children realized dorsals in complex onsets.

The fact that these children had a delayed acquisition does not have any negative implications for their grammar. These children are in the process of acquiring a grammar and the grammars of children with a delayed acquisition cannot be dismissed as being qualitatively different from the grammars of normally developing children (Barlow, 1997; Bishop, 1997; Morrisette et al., 2003; Pater and Barlow, 2003).

The data of the German children were collected in a picture naming task. The children were shown a picture and had to say the appropriate word. If they did not know the word immediately there were helped by an additional description. For example, they were shown a picture of a piano ([klaviːr]) and if they could not think of the word right away, they were told it has a lot of keys and is used to make music. If they still couldnt think of the word, they were told and asked to repeat the word. Of course, such repetitions were discarded from the data set.

In German, dorsals are allowed in complex onsets but not in simple ons (4).

(4) *German*

Tim, age 5;5

Word	*target*	*child's output*	*gloss*
Kette	**k**ɛtə	**t**ɛtə	'necklace'
Clown	**kl**aʊn	**kl**aʊn	'clown'

continue

Justin, age 4;6

Word	*target*	*child's output*	*gloss*
Gießkanne	**g**iːs**k**anə	**d**iːs**t**anə	'watering can'
Knochen	**kn**ɔ**x**ən	**gn**ɔ**t**ən	'bones'
Michelle, age 4;7			
Käse	**k**ɛːsə	**t**ɛːsə	'cheese'
Glocke	**gl**ɔ**k**ə	**gl**ɔ**t**ə	'bell'
Kleid	**kl**aɪt	**kl**aɪt	'dress'
Marvin, age 4;5			
Kuh	**k**uː	**t**uː	'cow'
Glocke	**gl**ɔ**k**ə	**gl**ɔ**t**ə	'bell'
Sebastian, age 6;6			
Külschrank	**k**yːlʃran**k**	**t**yːlʃran**t**	'refrigerator'
Knochen	**kn**ɔ**x**ən	**kn**ɔ**t**ən	'bones'

2.2. Dutch

The pattern in (4) is found in Dutch with a twist. dorsals are not allowed adjacent to stressed vowels. Since the dorsal in *klein* 'small' is not adjacent to a stressed vowel, a dorsal is allowed there, but the dorsal in *ku* 'cow' is adjacent to a stressed vowel, so it isn't allowed there. The data in (5) come from a three year old boy acquiring German, Dutch and English. Even though the pattern in (5) can be found in all three of his languages, the focus is on Dutch, since this is the language of interaction between him and his father, who collected the data. The data are all spontaneous speech.

(5) *Dorsals before unstressed sonorants, but not before stressed sonorants*

Tristan, age 3;0, Dutch

Word	*target*	*child's output*	*gloss*
kaal	kal	tal	'bald'
kikker	ˈkɪkəer	tɪtər	'frog'
Kermit	ˈkɛrmɪt	tɛrmɪt	'Kermit'
klein	kleɪn	kleɪn	'small'
kriebelen	kribələ	kribələ	'itch'
koppie krauw	ˈkɔpi kraʊw	tɔpi kraʊw	'head scratch'

2.3. English

In English the dorsals are distributed as in Dutch. Bills and Golston (2002) discuss and analyze data from a monolingual English girl acquiring central California English. The girl realizes dorsals only if they are followed directly by an unstressed sonorant (6). Dorsals may not appear immediately adjacent to a stressed vowel.

(6) *Dorsals before unstressed sonorants but not before stressed sonorants*
Sine, age 2;9–3;2, English, Bills and Golston (2002)

Word	*target child's*	*output*
candy	kændi	tændi
cow	kaʊ	taʊ
goats	gots	dots
glass	glæs	głæs
Chris	krɪs	kɰɻɪs
clue	klu	kłu
because	bəˈ**k**ɔz	pəˈ**t**ʌz
pancake	ˈpæˌ**k**e**k**	ˈpæˌ**t**e**k**
froggy	frɑ**g**i	fɰɻɑ**g**i
finger	fɪŋ**g**ər	fɪŋ**g**ʊ
triangle	traɪŋgł	ˈ**tsaɪ**ˌ**æŋgogogo**

In this section we have seen data of children acquiring three different languages, who all avoid dorsals immediately before a stressed vowel, but produce dorsals only before unstressed sonorant consonants (German), or before unstressed consonants and vowels (Dutch and English). In section 3 this patterns will be analyzed.

3. Analysis

We propose to account for the distribution of dorsals by means of two *positional markedness* constraints that conflict with a faithfulness constraint. Such a positional markedness constraint has been around in the literature in several guises, and has most recently been elaborated upon in (see Zoll, 2004, and references therein). Positional markedness is a theory about licensing in OT. Marked structures either occur or can not occur in certain positions.

In our case, the marked structure is the [dorsal] feature. It is marked crosslinguistically, in the sense that languages that have them, also have other places of articulation, but not vice versa. Another reason to assume it is marked is that it appears in onsets after labials and coronals (Levelt, 1994; Fikkert and Levelt, 2006). In other words, if children have a dorsal as an onset, they also have labial and coronal onsets.

The position in which it is allowed to occur is a branching, or complex onset and dorsals are prohibited to occur before stressed vowels. Onsets of unstressed syllables are usually seen as lenition contexts (Hayes, 1995). Flapping of coronals, for example, takes place in the onset of unstressed syllables. The onsets of stressed syllables, on the other hand, are usually seen as licensors of more contrast.

There are at least three arguments that show that complex onsets are independent sub-syllabic constituents. First, they are acquired later (Fikkert, 1994; Levelt and van de Vijver, 2004). Second, a language that has complex onsets also has simple onsets, but not vice versa (cf. Prince and Smolensky, 1993) and, third, there are restrictions on the distribution of complex onsets (Oostendorp, 1995). complex onsets in Dutch occur in the onset of full-voweled syllables and nowhere else.[4]

The fact that marked features have a limited distribution is not unheard of, of course. In Dutch lax vowels are marked and they occur in closed syllables only (Oostendorp, 1995).

Before defining the constraint, we have to briefly discuss an alternative to positional markedness and why that alternative is less suited to explain the distribution of dorsals. In the theory of *positional faithfulness* marked structures are limited to certain salient positions only if they are present in the input (Beckman, 1998; Lombardi, 1999). Salient positions are positions that play an important role psycholinguistically, such as the initial syllable, which plays an important role in processing and in lexical access (see references in Beckman, 1998). There is no evidence, as far as we know, to show that a complex onset is either more salient or less salient than a simple onset. Such evidence would give us a functional explanation of why neutralization of dorsals (with respect to the target language) would occur in simple onsets but not in complex onsets. Because of the absence of this evidence we have decided to account for the distribution of dorsals in terms of positional markedness rather than in terms of positional faithfulness.

There are two positional markedness constraints (7) and (8).

(7) COINCIDE([dorsal], branching onset): A dorsal belongs to a branching onset.

$\forall x(x \text{ is [dorsal]}) \rightarrow \exists y(y=\text{Branching onset} \wedge \text{COINCIDE}(x,y))$

Assess one mark for each value of x for which the above statement is false.

Any dorsal encountered outside of a branching onset violates (7). This constraint is in conflict with a faithfulness constraint (9).

(8) *KV́

No dorsal is allowed immediately before a stressed vowel.

In fact, the constraint in (8) allows dorsals before unstressed sonorants: before unstressed vowels and before (obviously) unstressed sonorants in complex onsets. These constraints conflict with a faithfulness constraint.

(9) IDENT(dorsal): Input and output have the same value for dorsal.

Let α be a segment in S_1, and β be any correspondent of α in S_2. If α is [dorsal], then β is [dorsal].

It should be noted that the set of outputs allowed by (7) is a subset of the set allowed by (8). Whenever (7) outranks (9) dorsals are only allowed in complex onsets, irrespective of the ranking of (8). Dorsals are allowed in complex onsets and before unstressed vowels if (8) outranks the other two constraints.

The tableau in (10) illustrates that the ranking COINCIDE([dorsal], branching onset) $\gg$ IDENT(dorsal) prevents dorsals from appearing in simple onsets.

(10) *Kuh*

/kuː/	COIN([dorsal], br. ons.)	*IDENT(dorsal)	*KV́
[kuː]	*		*
☞ [tuː]		*	

Any dorsal in a simple onset will violate COINCIDE([dorsal], branching onset) and since in the grammar of the child this is apparently more harmonic than faithfully producing the dorsal, the ranking of the constraints must be as in (10). Even if the child assumes as input *tu* the grammar would still output *tu*. This is illustrated in (11).

(11) *Tuh*

/tuː/	Coin([dorsal], br. ons.)	*Ident(dorsal)	*KV́
[kuː]	*	*	*
☞ [tuː]			

In short, the grammar in (10) and (11) neutralize the dorsal contrast in simple onsets. Now, let us turn to complex onsets.

(12) *Kleid*

/klaɪt/	Coin([dorsal], br. ons.)	*Ident(dorsal)	*KV́
☞ [klaɪt]			
[tlaɪt]		*	

The fact that neither the candidate *kleit* nor the candidate *tleit* violates Coincide([dorsal], branching onset) relegates the decision on which candidate is optimal to Ident(dorsal). This constraint prefers *kleit* over *tleit*, if the child assumes *kleit* as input. This assumption leads to a match between the child's output and the target output. She concludes, therefore, that the grammar which yields *kleit* as output is a better model of the data than a grammar which yields *tleit* as output. Therefore, the child decides on *kleit* as input.[5]

The constraints as ranked in (10) and (12) result in a distribution of dorsals such that they are allowed in complex onsets and prohibited in simple onsets.

The distribution of dorsals found in (5) and (6) is accounted for in a grammar in which *KV́ outranks Ident(dorsal) which, in turn, outranks Coincide([dorsal], branching onset).

(13) *A small frog and a cow; Dutch (klein, kikker, koe)*

a. *klein*

/klɛɪn/	*KV́	*Ident(dorsal)	Coin([dorsal], br. ons.)
☞ [klɛɪn]			
[tlɛɪn]		*	

b. *kikker*

/kɪkər/	*KV́	*Ident(dorsal)	Coin([dorsal], br. ons.)
[kɪkər]	*		*
☞ [tɪkər]		*	*
[tɪtər]		**	

c. *koe*

/kuː/	*KV́	*IDENT(dorsal)	COIN([dorsal], br. ons.)
[kuː]	*		*
☞ [tuː]		*	

The same ranking accounts for the English data.

(14) *clue, froggi, cow*

a. *clue*

/kluː/	*KV́	*IDENT(dorsal)	COIN([dorsal], br. ons.)
☞ [kɫuː]			
[tɫu]		*	

b. *frAgi*

/frɑgi/	*KV́	*IDENT(dorsal)	COIN([dorsal], br. ons.)
☞ [frɑgi]			*
[frɑti]		*	

c. *cow*

/kɑʊ/	*KV́	*IDENT(dorsal)	COIN([dorsal], br. ons.)
[kɑʊ]	*		*
☞ [tɑʊ]		*	

In our approach the distribution of dorsals is connected to complex onsets. In adult targets such a coupling is not present. A language either allows dorsals or not (15).

(15) *All or nothing constraints*

Constraint 1	≫	Constraint 2	Outcome
*DORS	≫	IDENT(dorsal)	No dorsals anywhere
IDENT(dorsal)	≫	*DORS	Dorsals everywhere

Obviously, an analysis of the distribution of dorsals as described in section (2) cannot be casted in terms of the constraints in (15). This raises the question whether constraints are universal and if they are not where do they come from? This will be the topic of section (4).

4. Typology and Child Language

In OT each grammar is one particular ranking of a set of universal constraints, and different grammars can be found by reranking constraints. The set of constraints is assumed to universal: all grammars have the same constraints. Each grammar is a particular instantiation of a ranking of all constraints. There is therefore a tight connection between typology and developing grammars (Levelt and van de Vijver, 2004). Each developing grammar is a particular ranking and it is expected to parallel an adult grammar. Distributions of linguistic objects, such as dorsals, are brought under the purview of grammar in OT (Prince and Smolensky, 1993; McCarthy, 2002). In other words, constraint ranking is responsible for distributions. In section 3 we have seen the rankings that take care of the distribution of dorsals described in section 2. Now it is expected that these rankings would correspond to an adult language, but, as far as we know, this is not the case.[6]

If the constraints in (15) would be the only ones to account for the typology of dorsals, there could be no language in which dorsals occur in a limited environment. If a dorsal is allowed, it means that IDENT(dorsal) $\gg$ *DORS and therefore dorsals are allowed anywhere. This immediately shows that such an analysis is inadequate to deal with the data in section (2).

So now we are faced with the situation in which typology and acquisition data diverge. The explanation of our data falls short typologically, and the typology of dorsals (the constraints in (15)) cannot be mapped onto our data. A solution is to assume that there are limited differences between child language and adult languages. There are more such differences. For example, the speech of many children exhibits *consonant harmony* (Smith, 1973; Levelt, 1994; Bernhardt and Stemberger, 1998; Fikkert and Levelt, 2006). The consonants of a word have to agree in place of articulation. Amahl, whose phonological development is analyzed in Smith (1973), produces *gɔk* for *sock*, where all consonants are dorsal. No adult language has a comparable process. The data presented in this paper show that there are, in addition to apparent child specific processes, also child specific distributions. Future research should clarify why and how child language may differ from the target. What is needed is a theory that tells us what possible differences between child language and adult language are, given that we know there are such differences (Pater, 1997, 2002; McCarthy, 2002; Levelt and van de Vijver, 2004; Fikkert and Levelt, 2006).

If children invent or construct constraints, for example as a reaction on a growing lexicon, finding differences may not be surprising. A scenario might be the following (Fikkert and Levelt, 2006). At first, the child has no dorsals in word-onsets at all. Such a grammar may be described with the aid of a constructed constraint $*_{Wd}[$(dorsal). Crucially this constraints outranks a faithfulness constraint. The only PoA contrast is between Labial and other. At this point there is no reason for the child to assume any underlying dorsals. Lexicon Optimization (Prince and Smolensky, 1993) will select the input without dorsals. This is illustrated in (16), in which we use the *tableau des tableaux* technique of Itô, Mester and Padgett (1995). This allows a comparison of various mappings of input on output. The most harmonic input-output pairing is the one in which the input is identical with the output.

(16) *Tableau des tableaux for output [tu]*

Input	*Output*	$*_{Wd}$[(dorsal)]	IDENT(dorsal)
/kuː/	[kuː]	*	
/kuː/	[tuː]		*
/tuː/	[kuː]	*	*
☞ /tuː/	☞ [tuː]		

In the next step in its acquisition the child starts producing complex onsets. Now dorsals are needed, perhaps to keep lexical neighbors from the target apart. To store them, the grammar needs to change. For example, by constructing a constraint such as COINCIDE([dorsal], branching onset). This will lead to a grammar in which dorsals can be stored, as is illustrated in (17). The best input for the output [kleɪn] is one in which the underlying form is /kleɪn/. The underlying form /tleɪn/ results in the wrong output.

(17) *Tableau des tableaux for output [kleɪn]*

Input	*Output*	COIN([dorsal], br. ons.)	IDENT(dorsal)
☞ /kleɪt/	☞ [kleɪt]		
/kleɪt/	[tleɪt]		*
/tleɪt/	[kleɪt]		*
/tleɪt/	☹ [tleɪt]		

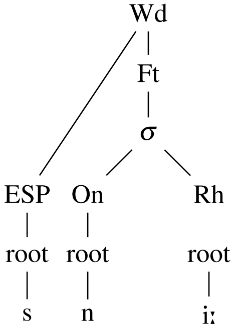
Funny constraints, such as the ones proposed here, must be grounded in a theory which explains the interaction between growth of the lexicon and the growth of the grammar. The theory must also explain what will happen to these constraints: will they recede in oblivion after they have done their useful work, or will they rear their heads in certain situations. Given the architecture of OT one would expect the latter.

4.1. Further issues

It might be argued that learning the distribution of dorsals is a matter of learning to distinguish the correct cues in the correct environment (Vanderweide, 2005). The cues for dorsals before sonorants may be stronger than the cues for dorsals before vowels. There are two problems with such an approach. The first is that the cues for dorsals (a burst to indicate stops and formant transitions to indicate PoA) are as strong before vowels as they are before liquids. From the point of view of cue-based learning the distribution of dorsals discussed in this paper cannot be explained. The second problem is that even if cues were responsible for the observed patterning of dorsals, it would still not explain the typological lack of such a pattern.

Another possible line of analysis of these data is to say that the children have analyzed these clusters as complex segments, rather than as complex onsets. However, data from Tristan show that he is really getting to grips with Dutch syllable structure. He realizes target dorsal-nasal clusters, such as in [kniː] as [sniː]. An analysis might be that the stop and the nasal are to close in sonority to make up a good cluster in Tristans grammar. Yet, he prefers to realize both root nodes. In order to do, he realizes the dorsal outside of the syllable. There it must be realized as [s], in accordance with the grammar of Dutch (Hulst, 1984).

(18) *[sniː]*



ESP = Extrasyllabic Position.

This shows that Tristan tries to fit dorsals in his syllable grammar, and that he does not entertain the possibility that they are part of complex segments. Other German children also provide arguments against such an analysis (Ott, 2004; Höhle et al., 2006). These children realize fricatives as stops. They prefer to realize a word as [ʃlɑŋə] 'snake' as [tlɑŋə] and never as *[klɑŋə]. Again, this shows that these children are working on their syllable structures and do not pursue the hypothesis that such clusters are actually complex segments. Another issue is the following. It might be said that languages such as German, Dutch and English lack complex onsets in which the consonants have the same PoA. There are no onsets such as *tl* or *pw* (see Booij, 1995; Wiese, 1996; Hammond, 1999, for Dutch, German and English respectively). So, the children might not be able to avoid complex onsets with dorsals (by producing complex onsets in which there is one PoA feature, such as *tl* or *pw*) even if they do not yet produce dorsals in simple onsets. There should be a higher ranked markedness constraint against clusters in which the consonants have the same PoA. It is, however, both from the point of view of the target language and from the point of view of acquisition not so clear what the nature of such a constraint is. In all of these languages borrowed words that start with *tl* are realized unaltered. A word like *Tlingint*, the name of an Inuit people, is pronounced [tlɪŋgɪt] in German (Wiese, 1996), Dutch and English.[7] If there were really a constraint against *tl* clusters one would expect that the output of the borrowing would somehow have been adjusted. This does, of course, not mean that we claim that complex onsets such as *tl* or *pw* are fine. They are not. But it is not entirely clear how to account for their apparently not ungrammatical strangeness.

In acquisition there are indications that such clusters are not avoided, but rather play a crucial role in acquisition. One piece of evidence has been adduced above, where some children realize [ʃlɑŋə] 'snake' as [tlɑŋə]. Another piece of evidence comes from Fikkert (1994), who claims that many Dutch children go through a stage in which target *kl* is realized as *tl*.[8]

5. Conclusions

In this paper we have shown that in child language dorsals may either be present in complex onsets but not in simple onsets or be present in complex onsets and in simple onsets before an unstressed vowel. This is explained by means of positional markedness constraints. One of them allows dorsals in complex onsets and bans them from other constituents and the other one bans them before stressed vowels. In combination with a constraint which requires a faithful mapping between input and output dorsals, the patterns can be explained.

We have also shown that the standard theory on the typology of complex onsets and the distribution of dorsals (15) does not describe our data. As a consequence, our analysis makes incorrect typological predictions–there should be adult languages in which dorsals are present in complex but not in simple onsets, but there are no such languages. Our data further corroborate the idea that there is substantive varation between child languages and adult targets and we have pointed out that a theory explaining why and how child language differs from adult targets is needed.

Notes

1. We would like to thank: Paula Fikkert, Joe Pater, Curt Rice, Hubert Truckenbrodt, Jeroen van der Weijer, Susanne Winkler and the audiences of the SFB-Kolloquium in Tübingen, the workshop on language acquisition in Konstanz and the audience of OCP4 at Rhodes, and all the children.
2. The framework of this paper is Optimality Theory, but other theories of language would face similar conclusions. In the Principles and Parameters approach, for example, the principles and parameters are universal, leaving no room for child specific properties of grammar.
3. The other 6 children replaced fricatives by stops. Their pattern is irrelevant for the purpose of this paper and will not be discussed here. See Höhle, van de Vijver and Ott (2006) for an analysis.
4. Fikkert (1994); Fikkert and Levelt (2006) note that dorsals are allowed in codas, even at very early stages of acquisition. We have not enough data in our sample to evaluate whether this is also true of the grammars of the children discussed in this paper. In principle a constraint DORSAL], which states that codas are dorsal would describe such a pattern. It should be ranked above the constraint in (7).
5. The decision on the nature of the final segment of the input (voiced or voiceless) is irrelevant in this context.
6. At the featural level there are languages the allow complex segments, but not their simple counterparts Prince and Smolensky (1993).
7. Kudos to Richard Wiese for bringing this point to our attention. The Dutch realization is the one of the first author, and confirmed by native speakers of Dutch. The English pronunciation comes from Kirsten Brock and is confirmed by native speakers.
8. Manuela Schröder p.c., who is working on her MA-thesis, tells us that she found a similar stage for German children acquiring clusters with dorsals.

References

Barlow, J. A.
1997 *A Constraint-based Account of Syllable Onsets: Evidence From Developing Systems*. Ph. D. thesis, Indiana University.

Beckman, J.
1998 *Positional Faithfulness*. Ph. D. thesis, University Of Massachusetts, Amherst.

Bernhardt, B. H. and Stemberger, J. P.
1998 *Handbook of Phonological Development: From the Perspective of Constrained-based Nonlinear Phonology*. San Diego, CA: Academic Press.

Bills, S. and Golston, C.
2002 Prosodic licensing and linear licensing in English acquisition. Ms. California State University Fresno.

Bishop, D. V. M.
1997 *Uncommon Understanding. Development and Disorders of Language Comprehension in Children*. East Sussex, UK: Psychology Press.

Booij, G.
1995 *The Phonology of Dutch*. Oxford: Clarendon Press.

Fikkert, P.
1994 *On the Acquisition of Prosodic Structure*. Ph. D. thesis, Holland Institute of generative Grammar, University of Leiden.

Fikkert, P. and Levelt, C.
2006 How does place fall into place: The lexicon and emergent constraints in children's developing phonological grammar. Ms. Radboud University Nijmegen and Leiden University.

Gnanadesikan, A.
2004 Markedness and faithfulness constraints in child phonology. In *Constraints in Phonological Acquisition*, R. Kager, J. Pater, and W. Zonneveld (eds.), 73–108. Cambridge University Press.

Hammond, M.
1999 *The Phonology of English*. Oxford: Clarendon Press.

Hayes, B.
1995 *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.

Höhle, B., van de Vijver, R., and Ott, S.
2006 The effect of phonotactic constraints in German-speaking children with delayed phonological acquisition: Evidence from production of word-initial consonant clusters. *Advances in Speech-Language Pathology* 8(4): 323–334.

Hulst, H. v. d.
1984 *Syllable Structure and Stress in Dutch*. Dordrecht: Foris.

Itô, J., Mester, A., and Padgett, J.
1995 Licensing and underspecification in optimality theory. *Linguistic Inquiry* 26: 571–614.

Levelt, C. C.
1994 *On the Acquisition of Place*. Ph. D. thesis, Holland Institute of generative Grammar, University of Leiden.

Levelt, C. C. and van de Vijver, R.
2004 Syllable types in cross-linguistic and developmental grammars. In *Constraints in Phonological Acquisition*, R. Kager, J. Pater, and W. Zonneveld (eds.), 204–218. Cambridge University Press.

Lombardi, L.
1999 Positional faithfulness and voicing assimilation in optimality theory. *Natural Language and Linguistic Theory* 17: 267–302.

McCarthy, J.
2002 *A Thematic Guide to Optimality Theory*. Cambridge, UK: Cambridge University Press.

Morrisette, M. L., Dinnsen, D. A., and Gierut, J. A.
2003 Markedness and context effects in the acquisition of place features. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 48(3/4): 329–355.

Oostendorp, M. v.
1995 *Vowel Quality and Syllable Projection*. Ph. D. thesis, Tilburg University.

Ott, S.
2004 *Der Einfluss des phonotaktischen Wissens bei kindlichen Aussprachestörungen. Eine Analyse der Produktionen von Konsonantenclustern bei alveolarisierenden und plosivierenden Kindern*. Master's Thesis.

Pater, J.
1997 Minimal violation and phonological development. *Language Acquisition* 6: 201–253.

Pater, J.
2002 Form and substance in phonological development. In *WCCFL*, volume 21, L. Mikkelsen and C. Potts (eds.), 348–372. Sommerville, MA. Cascadilla Press.

Pater, J. and Barlow, J. A.
2003 Constraint conflict in cluster reduction. *Journal of Child Language* 30: 487–526.

Prince, A. and Smolensky, P.
1993/2004 Optimality theory: constraint interaction in generative grammar. Ms. Rutgers University, New Brunswick and University of Colorado at

Boulder [Technical Report No. 2, Rutgers University Center for Cognitive Science.] Malden, MA, Blackwell.

Smith, N. V.
1973 *The Acquisition of Phonology*. London: Cambridge University Press.

Vanderweide, T.
2005 *Cue-Based Learning and the Acquisition of Pre-Vocalic Clusters*. Ph. D. thesis, University of Calgary.

Wiese, R.
1996 *The Phonology of German*. Oxford: Clarendon Press.

Zoll, C.
2004 Positional asymmetries and licensing. In *Optimality Theory in Phonology: A Reader*, J. McCarthy (ed.), 365–378. Blackwell.

Phonological knowledge in compensation for native and non-native assimilation

Isabelle Darcy, Franck Ramus, Anne Christophe, Katherine Kinzler & Emmanuel Dupoux

Abstract

We investigated whether compensation for phonological assimilation in the first language depends on language specific knowledge of phonological processes. To this end, we tested two different assimilation rules, one that exists in English and involves place of articulation, and another that exists in French and involves voicing. Both contrasts were tested on speakers of French and American English. In two experiments using a word detection task, we observed that participants showed a significantly higher degree of compensation for phonological changes that correspond to rules existing in their language than to rules that do not exist in their language (even though they are phonologically possible since they exist in another language). Thus, French participants compensated more for voicing than place assimilation, while American English participants compensated more for place than voicing assimilation. In both experiments, we also found that the non-native rule induced a very small but significant compensation effect, suggesting that both a language-specific and a language-independent mechanism are at play. Control experiments ensured that changes in stimuli were clearly perceived in isolation, compensation then being due to the phonological context of change, rather than to specific phonetic cues. The results are discussed in light of current models of lexical access and phonological processing.

1. Introduction

Understanding how words are recognized in continuous speech presents a particular challenge because the acoustic and phonetic shape of a word may be severely distorted in continuous speech compared to when that word is spoken in isolation. Words in sentences can be up to twice as short as words spoken in citation form. This higher speaking rate results in a number of acoustic changes due to co-articulation between the segments within and between the words (Church 1987; Liberman, Cooper, Shankweiler, and Studdert-Kennedy 1967; Trubetzkoy 1958). Even more dramatically, some language-specific phonological rules substitute, insert or delete entire segments as a function of speaking rate or phonological context (see Table 1). Such changes can potentially disrupt lexical recognition, since they can neutralize existing contrasts between phonemes, and hence contrasts between lexical items. In English, for example, place assimilation affects coronal stops, which take on the place of articulation of the following stop in connected speech (Barry 1992; Ellis and Hardcastle 2002; Nolan 1992). Hence the compound *football* may be realized as *foo[p]ball.* In French, voicing (glottal) assimilation voices obstruents before voiced obstruents, and devoices them before unvoiced obstruents (Dell 1995; Féry 2003; Wetzels and Mascaró 2001; Snoeren, Hallé and Segui 2006). So, the same word *football* tends to be realized as *foo[d]ball.* Such rules are common across the world's languages and tend to be productive, applying systematically to novel items. Moreover, when several rules coexist in a language, they can be chained to one another, resulting in large changes in surface word forms. For instance in French, the rules of nasal-obstruent simplification or word-final liquid deletion (Casagrande 1984; Dell 1995; Féry 2003) can be chained with regressive glottal assimilation: the sequence *table carrée* [tabl+kaʁe] 'square table' can thus become [tapkaʁe] in casual speech.

Even though there is considerable debate in the phonetic literature as to whether the phonetic change is complete or leaves traces of the original segment (Ellis and Hardcastle 2002; Féry 2003; Féry et al., this volume; Nolan 1992), it remains true that these rules substantially affect the phonetic shape of words. This in turn may render the identification of lexical entries problematic. The surprising fact is that these phonological changes seem to matter very little in everyday continuous speech recognition. In fact, most people are not even aware of the existence of these phonological changes. This calls for an explanation. What are the mechanisms responsible for robust lexical access despite near neutralizing changes induced by phonological rules?

Table 1: Examples of phonological rules which change the shape of words according to phonological context.

Language, Type of rule	Source	Featural Rule Description	Example
French, Regressive devoicing	Dell 1995; Féry (2003)	$\text{Obs}_{[+vd]} \rightarrow \text{Obs}_{[-vd]}$ / _ (#)$\text{Obs}_{[-vd]}$	*robe sale* ro/**b**#s/ale ‘dirty dress’ → [ʀɔpsal]
Dutch, Progressive devoicing	Wetzels and Mascaro (2001)	$\text{Obs}_{[+vd]} \rightarrow \text{Obs}_{[-vd]}$ $\text{Obs}_{[-vd]}$(+) _	/v/allen ‘to fall’ o/p+**v**/allen ‘to strike’ → o[pf]allen
English, Regressive place	Wells (1982: 55)	$C_{[cor]} \rightarrow C_{[vel]}$ / _ (#)$C_{[vel]}$	*good girl* goo/**d**#g/irl → [gʊggɜːl]
German, Progressive place	Wiese (1996)	$C_{[cor]} \rightarrow C_{[lab]}$ / $C_{[lab]}$ _	*geben* ‘to give’ /gebən/ → [gebm̩] *halten* ‘to hold’ /haltən/ → [haltn̩]
Turkish, [back] Harmony	Roca and Johnson (1999: 154)	$V \rightarrow V_{[+\text{ back}]}$ / $V_{[+\text{ back}]}$ C_0 + C_0 _ C_0#	[ip+in] ‘rope’ (Gen.sg.) ; [sap+ɯn] ‘stalk’ (Gen.sg.)
French, Liquid deletion	Dell (1995)	Liq → Ø / Obs_#Obs	*table jaune* ‘yellow table’ /tab**l**#ʒon/ → [tabʒon]
French, Nasal- obstruent simplification	Féry (2003)	ṼObs → ṼN / _(#)N	*Langue maternelle* ‘native language’ lan/g#m/aternelle → lan[ŋm]
English, r - insertion	Wells (1982: 58)	Ø → r / V(#)__ (#)V [-high]	*sofa is* /sofa#iz/ → [sofəɹis]

We review three classes of mechanisms that have been proposed in the literature. We call them *lexical compensation*, *phonetic compensation* and *language-specific phonological inference*. Models presented within a class are not assumed to be interchangeable, and the grouping of models into classes is based on predictions models make regarding three crucial features of compensation. The purpose of this paper is not primarily to distinguish between processing architectures or modeling details (which would require many more experiments), but rather to understand more in depth some aspects of compensation, given contradictory evidence in the literature.

1.1. Lexical Compensation

The first class of compensation mechanisms uses lexical knowledge. Since we know the words in our language, we can match the incoming signal with our stored list and pick the closest and/or most likely candidate available. This strategy essentially treats phonetic variation as random noise, and uses lexical and higher-order context to recover the signal from that noise. It is actually put to use in several speech recognition systems, and their mere existence attests the feasibility of such a mechanism. There is some evidence in psycholinguistics that lexical access incorporates robust mechanisms that resist input degradation. For instance, in running speech, lexical recognition is resistant to mispronunciations; participants might even have a difficult time to detect mispronunciations in fluent speech (Marslen-Wilson and Welsh 1978), and 'hallucinate' phonemes replaced by noise on the basis of lexical (and phonetic) proximity (Samuel 1981, 1996, 2001). Recent models of lexical recognition have implemented such robustness by relying on multiple activation of lexical candidates and competition between them (see the *Cohort* model, Marslen-Wilson and Welsh 1978; the *TRACE* model, McClelland and Elman 1986; and *Shortlist*, Norris 1994). This insures that whenever a degraded input is presented, several lexical candidates will be activated. Lexical competition, plus potentially higher-order expectations, ensures that the most plausible candidate is finally selected (Gow and Gordon 1995).

Although mechanisms like phoneme restoration may account for part of phonological compensation effects, they fail to distinguish between one-feature mispronunciations (which are often noticed) and one-feature assimilations (which are hardly ever noticed). Lahiri and Marslen-Wilson (1991, 1992) therefore developed a model of compensation based on underspecification theory (Archangeli 1988; Kiparsky 1985; Pulleyblank 1988), which explicitly implements regular phonological variation within lexical representations: They assume featurally "underspecified" lexical representations for words (FUL, Featurally Underspecified Lexicon, see Lahiri and Reetz 2002), for precisely those features that display regular variation. For instance, in English coronal stops would be unspecified for place, whereas labial or velar stops would be specified for place. Words containing coronal stops would thus have a gap in their featural specification; as a consequence, a deviant phonetic input could be mapped onto an unspecified segmental slot. Therefore, even if the sensory input differed in one position by one feature, its representation could nevertheless activate the appropriate lexical entry (see also Marslen-Wilson, Nix, and Gaskell 1995). This theory predicts an asymmetry in the recognition of lexical items depending on whether or not they contain unspecified segments. Using cross-modal priming, Lahiri and Reetz found that the deviant nonword stimulus **Bah[m]* triggered as much priming for the target *Zug* ('train' semantically related to *Bahn*) as the unchanged word *Bahn* 'railway' (where the coronal /n/ is assumed to be unspecified for place). In contrast, and consistent with their prediction, the deviant stimulus **Lär[n]* did not prime the target *Krach* 'bang', whereas the unchanged word *Lär[m]* 'noise' did (/m/ being specified as labial, only labials could map onto this slot). Note however that this result was not replicated by Gow (2001) who found equal priming for two similar conditions in English. Although the underspecification model cannot be fully equated with other models of lexical compensation, the predictions of all these models are similar.

Lexical compensation mechanisms have two crucial features. First, they rely on stored lexical items, and hence only work for restoring the phonological shape of actual words – not nonwords. Second, in their rudimentary form, they are insensitive to phonological context: the best-matching lexical item is selected based on the local phonetic cues and optionally the semantic and/or syntactic context. Crucially for the present experiments, the activation and selection of the most appropriate lexical item does not take into account the phonological context in which the changes occur, and whether these changes are systematic in the language or not.

Regarding the first feature (compensation for nonwords), most studies have used real words to assess compensation for assimilation. Using phoneme detection though, Gaskell and Marslen-Wilson (1998) found results with nonwords that were parallel to those of real words, although the amplitude of the effect was smaller. This effect on nonwords is impossible to account for with lexical compensation and suggests that compensation for assimilation is at least partly due to a non-lexical mechanism (see also Gaskell, Hare, and Marslen-Wilson 1995; Mitterer and Blomert 2003; Mitterer, Csépe, and Blomert 2003; Weber 2001, 2002).

Regarding the second feature (sensitivity to context), there is some robust evidence that compensation is sensitive to the segmental context in which the change occurs. For instance, Gaskell and Marslen-Wilson (1996) used cross-modal priming to examine compensation for place assimilation in English and observed more priming when the context was viable (leam#bacon → LEAN) than when it was unviable (leam#gammon → LEAN). These results were replicated and extended using other methods and assimilation processes by Coenen, Zwitserlood and Bölte (2001), as well as by Mitterer and colleagues (Mitterer and Blomert 2003; Mitterer, Csépe, and Blomert 2003). This sensitivity to context is not predicted by the compensation model based on underspecification (see above), where **Bah[m]* is expected to be recognized as a token of *Bahn*, without any influence of the context. In sum, it seems that one crucial property of lexical compensation mechanisms, i.e. insensitivity to phonological context, does not hold for phonological compensation.

A third crucial feature is related to the language-specificity of context sensitive compensation. Indeed, the use of context for compensation could originate in sensitivity to perceptual salience which would be different across phonetic contexts. This possibility predicts that context effects and compensation are to be found also for processes that don't exist in the language, as long as the appropriate context is given. Alternatively, context effects could also reflect the application of a kind of phonological knowledge, e.g. a familiarity with a particular type of modification (language-specific knowledge of the processes at work in a given language). This option limits compensation phenomena and context effects to those processes that exist in a language. Contradictory results in the literature mirror a vivid debate as to whether compensation reflects language-specific knowledge or not. This third crucial feature is exactly the point of divergence between the two remaining classes of models.

Let us review first some of the evidence in favor of phonetic compensation, which is not dependent on language-specific processing, but rather takes place at a lower level of processing.

1.2. Phonetic compensation

This class of compensation mechanisms is based on acoustic/phonetic processes. The idea is to deal with compensation for phonological variation using those mechanisms that compensate for phonetic variation or coarticulation. Several decades of research in acoustic/phonetics have shown that acoustic cues relevant to a given segment are temporally spread out across adjacent positions (Bailey and Summerfield 1980; Stevens 1998). It has also been shown that the perceptual apparatus of listeners integrates multiple cues to the same feature (Best, Morrongiello, and Robson 1981; Hodgson and Miller 1996; Parker, Diehl, and Kluender 1986; Repp 1982; Sinnott and Saporita 2000; Summerfield and Haggard 1977; Treiman 1999). These effects seem to hold across languages, and might even not be specific to humans, since compensation for coarticulation has been observed in birds (e.g. Lotto, Kluender, and Holt 1997).

Gow (2001, 2002a, 2003) proposed a language independent processing mechanism called *Feature Cue Parsing* to handle both coarticulation and systematic phonological variation. In this mechanism, temporally distributed acoustic cues of feature values are grouped and integrated into segmentally aligned phonetic features (see also Fowler 1996; Fowler and Brown 2000). Gow's specific proposal is that feature parsing can account both for coarticulatory compensation and compensation for phonological assimilation, at least in the (frequent) cases where assimilation is not complete. Indeed, in most cases, the target phoneme contains phonetic traces or partial cues of the original unassimilated form (Ellis and Hardcastle 2002; Nolan 1992). The principle of feature parsing is the following: Complex segments that simultaneously encode two places of articulation are parsed onto two adjacent segmental positions, when the following context may attract one of the features. Attraction may take place when the following segment shares the same place of articulation as one of the two encoded in the preceding segment (Gow and Zoll 2002: 58, example 2). As a result, feature parsing may suffice to give an account of compensation for phonological rules, because the information used to parse the input is provided by the phonetic signal alone. For this same reason, this process is assumed to be language-independent. Supporting evidence is found in Gow (2001, exp. 1), where one existing process (place assimilation from coronals to labials, e.g. *green* becoming [grim]) was tested against a non-existing one (place assimilation from labials to coronals, e.g. *glum* becoming [glun]). No effect linked to experience with a given phonological assimilation process emerged (same priming effect in a lexical decision task, see also Gumnior, Zwitserlood and Bölte 2005, for asimilar lack of context effect in German).

Note that although Feature Parsing may work when assimilation is incomplete, it does not provide an appropriate explanation when assimilation is complete: in this case, articulatory features are not spread across adjacent segments. Yet, several experiments have shown that compensation does occur with tokens that were deliberately produced with complete assimilation of the target phoneme (Coenen, Zwitserlood and Bölte 2001; Gaskell and Marslen-Wilson 1996, 1998; Mitterer and Blomert 2003). Further, Nolan (1992) and Ellis and Hardcastle (2002) demonstrated that a substantial proportion of spontaneous place assimilatory changes in English seem to be complete: that is, they left no detectable acoustic traces of the underlying phoneme. In addition, Feature Parsing would have trouble handling cases in which assimilation apparently skips over 'transparent' consonants, like [m] in the Russian phrase /iz#mtsenska/ [is#mtsenska] 'from Mcensk' (Hayes 1984, Jakobson 1956). Similarly, cases where listeners are confronted to elision, insertion or a combination of several processes would be hard to explain. Thus, although the Feature Parsing model could account for cases of partial assimilation, it does not seem to be powerful enough or abstract enough, to deal with the full spectrum of phonological variation.

1.3. Language-specific phonological inference

A third class of mechanisms has been proposed to deal specifically with phonological sources of variation: phonological inference. This was first developed in Marslen-Wilson, Nix and Gaskell (1995). Basically, phonological inference would be a language-specific mechanism that undoes the effect of assimilation rules that apply during phonological planning in production. Whether this is obtained through some kind of rule-based "reverse" phonology (Gaskell and Marslen-Wilson 1996, 1998, 2001), or through a statistically based recurrent connectionist model (Gaskell, Hare and Marslen-Wilson 1995; Gaskell 2003), the principle is the same (even though processing issues are quite different). Such language-specific phonological inference mechanisms can account for the experimental results found with complete assimilation tokens presented above (Coenen, Zwitserlood and Bölte 2001; Gaskell and Marslen-Wilson 1996, 1998). Crucially, they also predict that the pattern of compensation should depend on the listener's language.

Several studies have been investigating the perception of assimilated forms in a variety of languages, such as English (Gaskell and Marslen-Wilson 1996, 1998), Dutch (Koster 1987 ; Quené, van Rossum and van Wijck 1998), Japanese (Otake, Yoneyama, Cutler, and van der Lugt 1996), German (Coenen, Zwitserlood and Bölte 2001; Weber 2001), Hungarian (Mitterer, Csépe and Blomert 2003) and French (Hallé, Chéreau, and Segui 2000; Rigault 1967; Snoeren, Hallé and Segui 2006). Up to now, a few of them (Mitterer, Csépe and Blomert; Otake et al.; Weber) present evidence in favor of such language-specific effects. However, they include a cross linguistic design in which listeners are presented with non-native phonology or ill-formed sequences. These results are therefore contingent on the problem of non-native speech perception and/or of phonotactic violations. In Mitterer, Csépe and Blomert (2003), Hungarian and Dutch listeners had to identify the Hungarian word /bal/ 'left', which can be realized with a final [r] (rather as [ba$^{\text{l}}_{\text{r}}$] with a complex articulation) when concatenated to the suffix [ro:l] 'from the' (i.e. [barro:l]), but only as [bal] before the suffix [na:l] 'at the'. Therefore, the realization [ba**r**na:l] is an inappropriate assimilation. The identification task involving compensation and access to a lexical representation produced context effects and language-specific effects: Hungarian listeners had an identification bias towards the canonical [bal]-form when hearing the viable assimilation [barro:l]. This bias was absent in Dutch listeners, who were unable to *identify* (i.e. to decide whether they hear [bal] or [bar]) the syllables in the viable context – without ([balro:l]), or with assimilation ([barro:l]). However, clear conclusions are difficult due to the fact that these non-native listeners are hearing both nonwords and non-native phonemes. This result could thus be due to a more difficult discrimination, as shown by the authors. Indeed, they found an important difference between identification and discrimination tasks. For both groups, discrimination is more difficult in viable, than unviable contexts, and showed no effect of native language, indicating that it might be performed on the basis of lower-level, universal representations. When engaged in identification tasks, Dutch listeners don't seem to make use of the phonetic information given through the complex articulation in the stimuli, which would enable them to compensate for the change as do Hungarians. The authors conclude that *identification* performance seems to be influenced by language-specific experience (Mitterer, Csépe and Blomert 2003: 2323).

Other cross-linguistic evidence comes from Otake et al. (1996), showing that Japanese, but not Dutch listeners, were able to use nasal place assimilation in Japanese words (e.g. in _to**nb**o_ ‘dragonfly’, where /n/ is realized as [m] vs. _ko**nt**o_ ‘tale’, with a dental [n]) to predict the post-assimilation context. This was the case despite the fact that the process tested (place assimilation in nasals) is present both in Japanese and in Dutch phonology (being optional in Dutch and obligatory in Japanese). Interestingly, Koster (1987) found that Dutch listeners were able to detect “a word ending in /n/” in assimilated [mb] sequences, but slower and with more errors than when it had no assimilation (*groe*[m] *boek*, vs. *groen book*). In this experiment (Koster 1987: 98 – 102), words were produced with “complete neutralization”, and half of the targets were having a lexical counterpart (*lijn – lijm* ‘line – glue’ are both words), half were not (*groen* ‘green’ but **groem*). For Dutch listeners, therefore, a change from [n] to [m] is neutralizing and potentially blurs a lexical distinction. In Japanese, moraic nasals are never contrasting with respect to place of articulation, there is no possible word **komto* in Japanese (only non moraic nasals are contrasting in place of articulation, _ta**ma**go_ ‘egg’ vs. _ta**nu**ki_ ‘rakoon’ or _ta**na**ko_ ‘tenant’). The difference in behavior between Dutch and Japanese listeners may be due to the fact that Dutch listeners are hearing both nonwords and a different phonetic system, while Japanese might show compensation because this kind of assimilation in Japanese is obligatory and therefore, the canonical underlying representation itself might reflect assimilation. Again, like for Mitterer, Csépe and Blomert (2003), conclusions are subject to the interpretation that Dutch listeners may not be able to perceive the moraic nasals in the same way as Japanese listeners do.

In Weber’s study (2001), phoneme monitoring for the German fricative /x/ was used to test whether non-native listening is influenced when the non-native input violates a native assimilation rule (fricative assimilation in German (*la*[x]*t* ‘laugh’ vs. *li*[ç]*t* ‘light’), being violated in Dutch nonword stimuli, e.g. [lixt]). Results showed that German, but not Dutch, listeners responded with a pop-out effect to violation of the German fricative assimilation rule. This effect is visible with non-native input though: the stimuli were recorded by a Dutch native speaker, and “sounded Dutch” (Weber 2001: 101). In experiment 3 and 4 of her experiments, the design avoided the problem of presenting non-native input, but stimuli still contained a violation in the domain of phonotactics, where assimilation is obligatory in German (fricative assimilation and regressive nasal assimilation within syllables). Her results are therefore not directly informative with respect to the processing of legal native sequences.

So far, evidence for language-specific listening has been obtained mainly through presenting non-native input to participants. In these conditions, such differences could also be due to violations of phonotactic constraints, or to unfamiliar sound categories, or even to syllable structure, in short, they are contingent on the problem of non-native speech perception. Therefore, the question remains, at least in the case of compensation for assimilation, whether processing of *legal* sequences in a *native* phonology is also dependent on phonological knowledge, or whether any change potentially reflecting assimilation would give rise to language-independent compensation effects (as suggested by Gow's results, 2001). In this sense, clear evidence in favor of language-specificity in processing native input is rather sparse. In sum, all these results indicate some language-specific elements in the processing of assimilated sequences, but do not give enough information about the way a possible model of word recognition would deal with assimilated words in a native language.

2. The present study

In order to further refine our understanding of language-specificity in compensation for assimilation, we designed a series of experiments, using a cross linguistic design but avoiding the problem of non-native speech perception. We included within the same language a native process as well as a non-native one, using exclusively the native categories of the listeners. We chose two comparable processes: regressive voicing and place assimilation. The first one exists in French, but not in English, whereas the second one exists in English, but not in French. Nevertheless both processes potentially neutralize phonemic contrasts of both languages. We therefore constructed French sentences containing occurrences of voicing assimilation (the native process) as well as occurrences of place assimilation (a non-native process). The same was done for English sentences.

In our experiments, listeners are processing only native speech, legal sequences and native phonetic categories in both conditions (place and voicing). Therefore any difference in compensation pattern that might emerge between the two conditions is hypothesized to reflect the use of language-specific knowledge of the process involved, rather than to differences arising from non-native speech processing.

As did most previous experiments on compensation for assimilation, we also considered context effects: occurrences of assimilation in our stimuli are either appropriate (i.e. surfacing in a suitable context for assimilation) or inappropriate (i.e. the context is normally not a trigger for the modification). Context effects are important because they show how the same sound can be interpreted differently when its phonological context is taken into account. We then distinguish two dimensions of modification in our stimuli: the native vs. non-native type of process, and within each, the appropriate (viable) vs. the inappropriate (unviable) context for the change. We also included a condition in which the target word surfaced without any change, to ensure that in this case, detection is robust. Table 2 summarizes these experimental conditions.

Table 2: Experimental conditions for each type of process (native vs. non-native). Examples given for English stimuli.

Condition	Place (native)	Voicing (non-native)
viable	we[p] pants	bla[g] glove
unviable	we[p] socks	bla[g] rag
no-change	wet shoes	black rug

The task we use is word detection: this is similar to identification, except that the actual response of the subject is a “similarity interpretation” rather than a “choice between two forms”: targets words are presented auditorily and followed by a sentence containing the target. But in the sentences, the targeted word surfaces either with a change (viable or not) or without any change (baseline). Participants are requested to press a button when they think that the target presented is the same in the sentence. A *yes* response then indicates that the word in the sentence is treated as a token of the target. A *no* response indicates that the change altering the word blocks its interpretation as a token of the target. This design then permits to obtain a measure of the degree of tolerance for modifications altering word forms. This is what we understand as *compensation*, i.e. when a change is compensated for, undone, in order to recover the “original/canonical” form of the word. If we see a difference in compensation between the native and the non-native type of change, this would be evidence in favor of the use of some knowledge of phonological processes during word recognition. In Experiment 1, French listeners are hearing French sentences, in Experiment 2, American English listeners hear American English sentences.

3. Experiment 1

3.1. Method

3.1.1. *Stimuli*

Thirty-two target items were selected. They were all monosyllabic French nouns, with a C(C)VC structure. The target items consisted of two sets of 16 items: the Voicing Set and the Place Set, that were matched in average frequency (Place: 4238; Voicing: 4837, $t(15) = -0.4$, $p > .1$) according to the Brulex Corpus (frequency per 100 millions, from Content, Mousty and Radeau 1990, see the complete list of items in Appendix I). In the Voicing Set, all items ended in a final obstruent that was voiced for half of the items, and unvoiced for the other half. Sixteen matched nonwords ([nw]) were constructed by switching the voicing feature of the final obstruents (e.g. *robe* /rɔb/ ‘dress’ - *rope* /rɔp/ [nw], or *lac* /lak/ ‘lake’ - *lague* /lag/ [nw]). In the Place Set, final consonants were all coronal; half were nasals and half were stops. Sixteen matched nonwords were obtained by a change in the place feature (12 towards labial, 4 towards velar) of the final consonant (e.g. *moine* /mwan/ ‘monk’ - *moime* /mwam/ [nw] or *guide* /gid/ ‘guide’ - *guibe* /gib/ [nw]). Each of the 32 target items was associated with a triplet of context words. In French, context words were always adjectives since the standard noun phrase has the shape ‘determiner noun adjective’. Each adjective in a triplet corresponded to one of the experimental conditions: *viable change*, *unviable change*, and *no-change*. For the viable change condition, the adjective’s initial consonant was an obstruent agreeing in voicing or in place with the nonword matched to the target item, depending on the item set (e.g. in the Voicing Set: *rope sale* /rɔ**ps**al/ ‘dirty dress$_{[nw]}$’[1]; in the Place Set: *moime bavard* /mwa**mb**avar/ ‘talkative monk$_{[nw]}$’, respectively. The adjectives in the unviable change and no-change conditions both started with a neutral consonant which was not involved in the relevant assimilation process. For the voicing set, this neutral consonant was always a sonorant (nasals and liquids, as well as the standard French uvular fricative [ʁ]), that does not trigger voicing assimilation in French. In the Place Set, this neutral consonant was a sonorant, a coronal or labiodental fricative, or the coronal stop [d]; none of these consonants is involved in place assimilation in English. In all 3 conditions of both the Voicing and the Place set, the association (pseudo)noun-adjective always yielded a legal consonant cluster in French and did not contain any violation of voicing or place assimilation.[2]

Finally, 3 sentence frames were constructed for each of the 32 target items. A sentence frame consisted in a sentence beginning and sentence ending, where each of the three (pseudo)noun-adjective combinations could be inserted and resulted in a plausible sentence (e.g. *Elle a mis sa* __ __ *aujourd'hui.* 'She put on her __ __ today.'). Globally, the sentence frames were matched in number of words and position of the insertion slots across the Voicing Set and the Place Set. No occurrence of violation of voicing or place agreement occurred in the frames neither. Combining the three conditions with the three sentence frames gave rise to 9 actual sentences associated to each item. This resulted in a total of 288 sentences.

For purposes of counterbalancing, we defined three experimental lists. In each list, all three conditions were present for each item, but in different sentence frames. The sentence frames were rotated across the three lists, so that across the experimental lists all three conditions appeared in all three sentence frames. Thirty additional filler sentences were constructed that were similar to the experimental sentences (same kind of alterations on the target involving one feature, same proportion of identical (1/3) and changed words (2/3)), and served as training (N=18), or distractors (N=12). Modifications involved voicing, manner and place contrasts at the end or beginning of target words, in order to drive participant's attention to the precise form of words (e.g. target "cube" [kyb], filler sentence containing "gube" [gyb]). Crucially, these filler sentences did not contain any case of assimilation in either viable or unviable context, so that the feedback provided here was unambiguous and could not influence later participant's responses on test sentences.

The 288 test, 12 distractor and 18 training sentences were recorded by the first author, a female native speaker of French.[3] The 32 target words for the experimental sentences and 30 targets for filler sentences were recorded by a male native speaker of French. They were digitized at 16 kHz and 16 bits on an OROSAU22 sound board, and edited using the sound preparation software CoolEdit and Praat. The onset of the carrier word and the onset of the following adjective were marked through digital labels.

3.1.2. *Procedure*

This experiment was run using the *Expe6* stimuli presentation program (Pallier, Dupoux, and Jeannin 1997). The experimental trials consisted in the presentation of the target item (male voice), followed after 500ms of silence by a sentence (female voice). Participants are requested to press a button when they think that the target presented is the same in the sentence, and refrain from pressing otherwise. This instruction – together with the specific training – was given in order to draw their attention on the detail of pronunciation of words, i.e. on the *form* of words and not to the mere presence or absence of a target word in the sentence. For the same reason only a few distractor sentences were included. This instruction was important in order to make participants understand that they have to be precise in their judgments and not only press *yes* if they recognized semantically the target word in the sentence. Otherwise, such minimal differences would have been at risk to be ignored in a word detection task. Several studies (McClelland and Elman 1986; Norris 1994) show that a word is still recognizable even if changes altered its canonical form. The degree of "recognizability" is inversely proportional to the word's frequency and neighborhood density. We therefore chose frequent monosyllabic words, in order to augment the importance of any minimal change affecting the word form. Participants are told to respond as quickly as possible, without waiting until the end of the sentence. They were allowed in total 3000ms after the word onset (in the sentence) to make their response. After that delay, the next trial is initiated. Reaction times (RT) were collected but our main measure is the word detection rate for each condition. Using reaction times as the main dependent variable in our experiments was difficult because they were collected only for "yes" responses. As a result, RT are calculated on the basis of a variable amount of yes responses in the different conditions, and would possibly fail to be a valid estimation of the average reaction times.

During the training phase (18 sentences), feedback was provided whenever the participants gave an incorrect response, that is, failed to detect the target word or incorrectly pressed a button for a non-target (the training sentences did not contain any occurrence of viable or unviable context). During the test phase, responses were collected without feedback. The test phase was split into three blocks of 36 trials that were constructed such that a given test item appeared only once within each block. A pause was inserted after each block to allow participants to rest and concentrate. Order of trials within each block was separately randomized for each participant. The experiment lasted 20 minutes. Instructions appeared on the computer screen, and were completed orally by the experimenter when needed.

3.1.3. *Participants*

Eighteen French native speakers (all grew up monolingually, having only limited and late experience with English) were tested on this experiment, individually and in a quiet room. There were 11 women and 7 men, all living in the Parisian area. They ranged in age from 19 to 28 years. None of them had previously taken part in a similar experiment, and none of them reported any history of hearing impairment. They were randomly assigned to one of the three experimental lists. They were paid for participation.

We expected participants to detect the target words in the no-change condition, and to reject them in the unviable change condition (in this sense, the logic of our experiment is similar to that of Gaskell and Marslen-Wilson 1996, 1998). The performance on these two conditions serves as comparison basis for evaluating the responses in the viable change condition. If participants fully compensate for the phonological rule, they should detect the target word to the same extent as in the no-change condition, despite the fact that the target underwent the same featural change as in the unviable change condition. If there is no compensation for the phonological rule, participants should respond like in the unviable change condition, that is, reject the changed word as a non-target.

3.2. Control task: forced-choice judgment on spliced-out target words

To ensure that the critical items' final consonants were unambiguously perceived as changed or unchanged, we first carried out a pretest in which we excised all target words out of the carrier sentences and presented them in isolation in a forced-choice categorization task. Words were presented auditorily and followed by a 3 s. silence, during which participants had to tick the consonant they heard on a response sheet. They always were given a choice between the original consonant and the assimilated one. For the word *robe* 'dress' for example, the choice was between [b] (unchanged) and [p] (underwent voice assimilation). A free cell allowed them to report any better matching sound, if needed. The entire procedure lasted about 18 minutes.

Eighteen French native speakers who did not participate in the other study were recruited to take part in this control experiment.

3.3. Results

We report first the results from the pretest, summarized in Table 3. Standard error (SE) is given in parentheses. Results include the whole data set (all items and participants).

Table 3: Different consonant judgment rate (%) across contrast type and condition for French stimuli (n=18).

	Consonant different from unchanged target (%):	
	Place (SE)	Voicing (SE)
viable change	92 (0.9)	95 (0.7)
unviable change	90 (1)	97 (0.5)
no-change	9 (2)	2 (0.2)

This table shows clearly that both change conditions yield in majority "different consonant" responses, there is no significant difference between both change conditions (an Analysis of Variance – henceforth ANOVA – with subjects as random variable, restricted to both change conditions for place and voicing together, yielded no effect of *condition* (F(1,17) = 0.2, p>.6). Items in the no-change condition are judged largely as having a "similar consonant" (to 91% and 98%). Globally, *contrast type* has no effect either (F(1,17)=4.2, p>.05).

For the word detection task, we checked whether some items triggered too many errors in the baseline conditions, namely the no-change and unviable change conditions. All items that yielded detection values higher than 50% in the unviable change condition (i.e. more than 50% false alarms) or less than 50% in the no-change condition (i.e. more than 50% misses) were excluded. In this experiment, only one voicing item (*badge*) was dropped.

The percent detection rate was subjected to two ANOVAs, one with participants, one with items as random variable. The by-subjects ANOVA had one between-subjects factor, *group* (counterbalancing factor, 1, 2 or 3) and two within-subject factors, *condition* (viable change, unviable change or no-change) and *contrast* (voicing or place). The by-items ANOVA had one between-item factor, *contrast* and one within-item factor, *condition*. We observed a main effect of *condition* (F_1[2,30] =635.8, p<.0001; F_2[2,58]=448, p<.0001), a main effect of *contrast* ($F_1[1,15]=63.8$, p<.0001; $F_2[1,29]=54$, p<.0001), as well as an interaction between these two factors ($F_1[2,30]=55.2$, p<.0001; $F_2[2,58]=37.1$, p<.0001), suggesting that the two item sets behaved differently across the three conditions. The *group* factor showed no main effect and did not interact with the other two factors. Similarly, the same analyses declaring the factor *blocks* (1, 2 or 3) instead of *group* revealed that there were no effects of *blocks* in subjects or items, suggesting that repeated presentation of the same word targets across different blocks did not cause any benefit or cost in processing. Mean detection rates are displayed in Figure 1 as a function of contrast and condition.

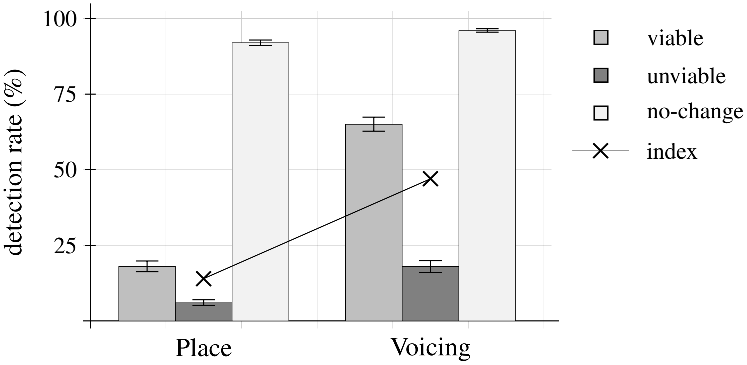


Figure 1: French listeners, French sentences: Detection rate in each condition, for both place and voicing assimilation types, N = 18.

Examination of mean detection rates revealed that the difference between the voicing and the place set was mainly in the viable change condition (65% for the voicing contrast vs. 18% for the place contrast, effect size 47%, $F_1[1,17]=72.4$, p<.0001; $F_2[1,29]=58.7$, p<.0001). In contrast, the other two conditions behaved similarly for both contrasts (14% vs. 06% in the unviable condition, effect size 8%, $F_1[1,17]= 2.1$, p>.1; $F_2[1,29]=2.9$, p=.094; 96% vs. 92% in the no-change condition, effect size 4%, $F_1[1,17]= 4.4$, p=.05; $F_2[1,29]=3.2$, p=.082).

Reaction times for this experiment are presented in table 4. The ANOVA analysis of mean reaction times restricted to the no-change and viable conditions for the voicing contrast[4], declaring the factors *group* (between-subject: 1, 2 or 3) and *condition* (within-subject: viable or no-change), revealed no effect of *group* (F[2,15]=0.44, p>.6), but a main effect of *condition*

(F[1,15]= 20.1, p<.0001). Participants responded slower to the viable change condition compared to the no-change condition. No significant interaction between both factors has been observed (F[2,15]=1.3, p>.2).

Table 4: French listeners, French sentences. Reaction times for each condition and each contrast.

Contrast	Condition	RT (ms.)	SD	
Place	Viable	1943	856	
Place	Unviable	2072	1023	
Place	No-change	1635	759	
Voice	Viable	**1672**	746	←
Voice	Unviable	1868	916	F[1,17]=19.2, p<.0001
Voice	No-change	**1566**	741	←

Mean times by subjects are comprised between 519 ms and 2107 ms (mean RT for n=18: 1582 ms). The experiment was fairly speeded: the time to make a response was limited, and participants should not wait until the end of the sentence. Overall, it should be noted that this experiment is demanding, speech rate is fast and contrasts are minimal. The slow RT we observed surely do not completely rule out the possibility of strategic responding. But we did our best to limit the risk of such a response pattern in our participants. A concern about offline strategic responding can however be reasonably rejected, as post-hoc analyses revealed no difference about the pattern of results according to slow vs. fast reaction times (ANOVA by subjects including the factor *RT* (fast vs. slow) and the factors *condition* and *type* revealed no interaction of the RT factor with both other factors).

To further refine our analysis, and to allow for a comparison of both sets with each other, we computed for each subject and item an index *x* of compensation (formula 1) on the basis of the number of yes-responses as a function of condition and contrast type (place vs. voicing). This index calculates the relative value of detection in the viable condition as a function of both other conditions. This allows obtaining the *ratio* of “viable” to “no-change”, controlling for response biases, or errors from the “unviable” condition.

$$(1) \qquad Compensation\ index = \frac{(detection_{viable\ \ change} - detection_{unviable\ \ change})}{(detection_{no\ \ change} - detection_{unviable\ \ change})}$$

The index x thus corresponds to the *degree* of compensation for either place or voicing type of change. If participants fully compensate for assimilation, they will detect the target word in the viable change condition as often as in the no-change condition: the index will be 1 (since the numerator and the denominator will be equal). If participants do not compensate at all for assimilation, they will respond to the target in the viable change condition as rarely as in the unviable change condition: the index will be 0 (since the numerator will be 0). Values of the index intermediate between zero and one will indicate partial compensation for assimilation.

We computed the compensation index for each participant and each contrast (mean index for participants is 0.65 (65%) for voicing and 0.14 (14%) for place), and used it as the dependent variable in an ANOVA with *contrast* as a within-subject (respectively between-items) factor. We found a significant effect of *contrast*, with a higher index of compensation for voicing than for place, confirming the fact that participants compensate significantly more for voice assimilation than place assimilation (65% vs. 14%, effect size 51%, $F_1[1,17]=77.4$, $p<.0001$; $F_2[1,29]=51.2$, $p<.0001$).

3.4. Discussion

Experiment 1 revealed two main results. First, French participants compensate for voicing assimilation in a context-sensitive fashion: viable contexts give rise to higher detection rates than unviable contexts. These results show a context effect comparable to the one observed by Gaskell and Marslen-Wilson (1998) with English listeners for a native assimilation process in English: place assimilation. We were also able to show that this compensation was not complete, however, since the compensation index only reached 65% (and was significantly different from 100%). This suggests that complete assimilation may not be the most natural case in French and that the word recognition processor is only able to compensate partially for such extreme cases. An alternative explanation could be that participants perform this recognition task integrating information from different processing levels simultaneously (*multiple readout hypothesis*, similar to Grainger and Jacobs 1996, or to the Race Model, Cutler and Norris 1979): the phonological level, representing a phonological form (recovered or not by a compensation mechanism), the lexical level, and a language independent phonetic level. A similar hypothesis (the dual task) has been evoked by Gaskell and Marslen-Wilson (1998), who observed that detection of phonemes in real words was higher than in nonwords. In our experiment, intermediate compensation (65%) may be the product of combining information from all levels: Faced with a (minimally deviant) word form, the lexical level leads to a "yes" response. The phonological level reinforces a "yes" response when the change is viable or has been compensated, whereas the phonetic form detector yields a "no" response.

The second main result from Experiment 1 is that French participants compensate much less for place assimilation, a rule that does not exist in French (the compensation index is only 14%), than for voicing assimilation. Since Gaskell and Marslen-Wilson (1998) previously obtained sizable compensation for place assimilation with British English participants and sentences (60% /t/-detection in assimilated *freigh[p b]earer*), this result corroborates that phonological compensation is language-specific. We will come back to this point in Experiment 2.

French participants nevertheless did compensate somewhat for place assimilation: even though the place change does not correspond to an existing rule in French, participants treated 18% of the words appearing in the viable change condition as tokens of the target as opposed to only 6% of the words in the unviable change condition ($p<.001$). The presence of a (small) context effect for this contrast (index value is 14%) suggests the existence of a language independent compensation mechanism in addition to the language-specific one; it nevertheless seems to be the case that the universal mechanism has a weak influence compared to the language-specific one, at least in a task involving complete changes. We are currently investigating whether this result reflects a general preference for homorganic consonant clusters, related for example to the high frequency of place assimilation phenomena across the world's languages.

So far, the difference observed in compensation between native and non-native assimilation suggests that compensation for assimilation reflects a phonological knowledge of these processes: This conclusion stems from the fact that French speakers showed greater compensation for voicing assimilation (a native rule), than for place assimilation (a non-native rule). However, this single experiment can not exclude the possibility that independent phonetic differences between voicing and place induced the results (see discussion section in Experiment 2). Indeed, it could be that voicing cues are intrinsically weaker than place cues in the context tested (VC#CV clusters), thus allowing for an easier acceptance of changed forms as being "the same", i.e. inducing more "compensation" before other obstruents which mask the preceding consonant. It could then happen that native listeners of other languages too would compensate more for voicing than place assimilation, whatever the rules actually present in their native language. At first sight, however, it seems not to be the case that voicing cues are intrinsically weaker than place cues. Indeed, voicing is a quite robust cue for several reasons: first, voicing is periodic in nature, distributed over lower regions of the spectrum than place, making it more robust to noise (Wright 2004). Second, because different acoustic parameters are involved (to name just a few: Vowel duration, duration of voiced portion in closure, closure duration, VOT-lag, F0) which all contribute to the voicing distinction (see Kohler 1984; Kingston and Diehl 1994, among others), listeners probably have more converging cues to this contrast. Indeed, place cues for stops are said to be weaker especially in this word-final cluster environment (VC#CV), where release burst is not reliable. Place cue markers are therefore restricted to VC-formant transitions, and are more variable in this VC position than in the CV position (Wright 2004; Jun 2004: 61). Because these are periodic as well, though, they resist quite well to masking, especially in optimal listening environments. An independent reason for considering voicing as being equal to place with regard to clarity is that the results of the control experiment did not show increased error rate for voicing items as compared to place, what would have been the case if voicing cues were less perceptible than place cues.

The possibility that place and voicing cues differ in strength in this environment seems implausible, and therefore we tend to interpret the results of the French listeners as support for a language specific compensation mechanism. However, in order to establish more strongly that compensation reflects language-specific knowledge of processes, and not only the language-independent use of phonetic properties, we need to test English participants with the same experimental design as we used for French participants. We expect the English participants to behave differently from the French participants: they should compensate more for place than for voicing assimilation. In contrast, if compensation for assimilation is largely language independent and based on differences between voicing and place, then English participants would behave much like French participants, and compensate more for voicing than for place assimilation.

English has no voicing assimilation rule, but a rule of place assimilation affecting coronal stops. Experiment 2 involves American English participants.

4. Experiment 2

4.1. Method

4.1.1. *Stimuli*

Following the same method used for French stimuli, 32 English words were selected as target items. They were all monosyllabic adjectives, with a C(C)V (C)C structure. Target items were split into two sets of 16 items: the Voicing Set and the Place Set. They did not differ in average frequency (per million, according to both the *Phondic* Database, and the Kucera and Francis Word Frequency as given in the MRC Psycholinguistic Database (Wilson 1988): voicing: 151 (K&F: 144), place: 156 (K&F: 152), t(15)=.06, p>.1; see the complete list of items in the appendix). In the Voicing Set, all items ended in a final obstruent, which was voiced for half of the items, and unvoiced for the other half. Sixteen matched nonwords ([nw]) were constructed by switching the voicing feature of the final obstruents (e.g. /nais/ (nice) - /naiz/ [nw], or /bik/ (big) - /bik/ [nw]). In the Place Set, all final consonants were coronals, and half were stops, half were nasals. Sixteen matched nonwords were obtained by a change in the place feature (towards labial or velar) of the final consonant (e.g. /swiːt/ (sweet) - /swiːk/ [nw] or /pleɪn/ (plain) - /pleɪm/ [nw]).

Each of the 32 target items was associated with a triplet of context words; In English context words were always nouns because the standard noun phrase in English is ‘determiner adjective noun’. Each noun in a triplet corresponded to one of the experimental conditions as defined in Experiment 1: viable change condition, unviable change condition, and no-change condition. For the viable change condition, adjectives started with an obstruent agreeing with the nonword matched to the target item; the nature of agreement was the same as described for Experiment 1 (place, e.g. [fæp pʌpi] ‘fat$_{\text{[nw]}}$ puppy’ or voicing, e.g. [blæg glʌv] ‘black$_{\text{[nw]}}$ glove’). Nouns in unviable change and no-change conditions for the Voicing Set started with a nasal or a liquid, consonants which are not involved in a voicing assimilation process. In the Place Set, nouns in both unviable change and no-change conditions started preferably with coronal sonorants, sometimes with coronal fricatives or the coronal stop [d] (the proportion of sonorants to obstruents is 5 to 3 in the place-stop list, and 2 to 6 in the place-nasal list). None of these consonants is involved in place assimilation processes in English. For the unviable change condition, the noun would be associated to the nonword matched with the target word (e.g. [blæg ɹæg] 'black$_{[\mathrm{nw}]}$ rag'). In the no-change condition, it would be associated to the target word itself (e.g. [blæk ɹʌg] 'black rug'). In all 3 conditions, the association (pseudo)adjectives-noun always yielded a legal cluster in English. There were no coronal-labial or coronal-velar clusters, in order to avoid spurious effects due to violation of the place assimilation rule.

Finally, 3 sentence frames were constructed for each of the 32 target items following the same method as used for French sentences. This resulted in a total of 288 sentences. Three experimental lists were defined similarly to those used in Experiment 1.

The 288 test, 12 distractor and 18 training sentences were recorded by the fourth author, a female native speaker of American English (her speech corresponding to General American standard), living in New Haven, CT. Target words were recorded by a male native speaker of American English from New York. They were digitized at 16 kHz and 16 bits on an OROSAU22 sound board, and edited using the sound preparation software CoolEdit and Praat. Onsets of the carrier words and onsets of the following adjectives were marked through digital labels.

4.1.2. *Procedure*

The same procedure was used for the presentation of the stimuli. However, we used the *E-prime* stimuli presentation program (www.pstnet.com/e-prime/-default.htm) instead of *Expe6*, due to hardware reasons. We also slightly modified the instructions: Participants had to press a "yes" button when they thought that the target was present in the sentence, and a "no" button otherwise.

4.1.3. *Participants*

Twenty-six Americans aged from 18 to 53, from the North-East of the U.S. (mainly New England), were tested on this experiment in Paris (France), in Providence (RI), New Haven (CT) and Amherst (MA). They all grew up monolingually, and came roughly from the triangle between Washington DC in the south, Chicago in the West and Boston in the North-East. None of them had previously taken part in a similar experiment and none of them reported any auditory deficits. They were paid for participating. All of them had late experience with French, 19 of them were living in France by the time of testing. They were tested on French sentences in the same testing session, half of them before American English, half of them afterwards. Nine participants were highly fluent in French; the 17 remaining were beginning learners. Their results on French sentences are presented in Darcy, Peperkamp and Dupoux (2007).

4.2. Control task: forced-choice judgment on spliced-out target words

As in Exp. 1, all target words were excised out of the carrier sentences and presented in isolation in a forced-choice categorization task. Sixteen American native speakers who did not participate in any of the previous studies were recruited to take part in this control experiment.

4.3. Results

Table 5 presents the results of the forced-choice categorization task. Results include the whole data set (all items and participants).

Table 5: Different consonant judgment rate (%) across contrast type and condition for American English stimuli (n=14).

	Consonant different from unchanged target (%):	
	Place (SD)	Voicing (SD)
viable change	74 (3)	78 (1)
unviable change	78 (2)	77 (1)
no-change	23 (4)	17 (3)

As can be seen from Table 5, both change conditions yield an equal amount of "different consonant" responses, there is no significant difference between both change conditions (an ANOVA with subjects as random variable, restricted to both change conditions for place and voicing together, yielded no effect of *condition* (F(1,13) = 2.3, p>.1). Items in the no-change condition are judged largely as having a "similar consonant" (to 80% on average). Globally, contrast type has no effect either (F(1,13)=0.1, p>.6).

One striking difference compared to the French results (see Table 3) is the higher error rate visible in the American English categorization results. However, this difference is not central to our argument. The most critical result to be seen in both control experiments is the absence of any difference in the "clarity of changes" between place and voicing targets, given the suggestion made above that voicing may have less clear cues, therefore favoring compensation over place targets. For both experiments, the answer is "no": in isolation, cues seem to be equal for voicing and place targets, and can not explain any observed differences in behavior. We return to the question of higher error rate in the discussion section for Experiment 2.

Using the same criterion for item rejection as in Experiment 1, 4 items were rejected, 1 in the Voicing set, 3 in the Place set.

Mean detection rate was subjected to two ANOVAs, one with participants, one with items as random variable. The participants ANOVA declares the between-subject factor *group* (1, 2 or 3), and two within-subjects factors: *contrast* (place vs. voicing) and *condition* (viable change vs. unviable change). As above, the by item ANOVA declared one between item factor *contrast* and one within-item factor, *condition.* In the participant analysis, no effects related to the factor *group* became visible. We observed a main effect of *condition* (F_1[2,46]=468.9, p<.0001; F_2[2,52]=181.9, p<.0001). The *contrast* effect was almost significant by participants, but not by items (F_1[1,23]=3.5, p=.07; F_2 [1,26]=0.3, p>.1). We found an interaction between these two factors that was significant only by participants, marginal by items (F_1[2,46]= 40.2, p<.0001; F_2[2,52]=2.7, p=.07), evidencing that they behave differently according to the contrast type (place vs. voicing) across conditions. Items display more variability, to which we will return below. Mean detection rates as a function of contrast and condition are displayed in Figure 3 (see below).

The viable change condition yielded 33% detection responses for the voicing contrast, and 46% for the place contrast, a significant difference by participants (effect size 13%, F_1[1,25]=32, p<.0001 ; F_2[1,26] =1.7, p>.1). The no-change condition was very similar in both contrasts (94% detection for place vs. 91% for voicing, effect size 3%, F_1[1,25]=1.8, p>.1 ; F_2[1,26]=0.6, p>.1). Detection rate in the unviable change condition was different between the place and the voicing contrast, significantly only by participants (11% vs. 18% for each contrast respectively, effect size 7%, F_1[1,25]=11.4, p<.01; F_2[1,26]=1.3, p>.1).

Reaction times for this experiment are presented in table 6. The analysis of mean reaction times restricted to the no-change and viable conditions for the place contrast, declaring the factors *group* (between-subject: 1, 2 or 3) and *condition* (within-subject: viable or no-change), revealed no effect of *group* (F[2,23]=1.2, p>.3), but a main effect of *condition* (F[2,46]=7.3, p<.002). Participants responded slower to the viable change condition compared to

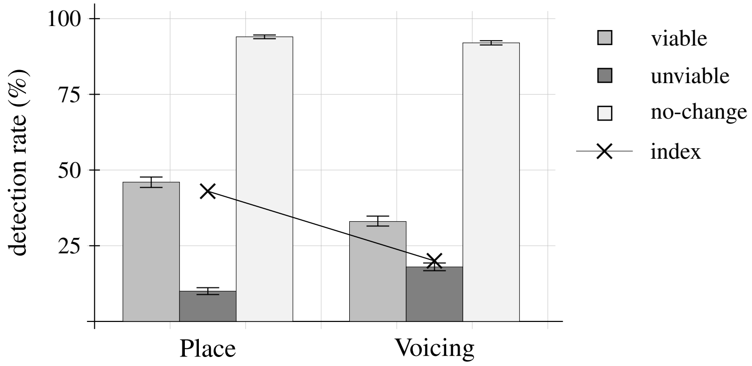


Figure 2: American listeners, American English sentences: Detection rate in each condition, for both place and voicing assimilation types, N = 26.

the no-change condition. No significant interaction between both factors has been observed.

Table 6: American listeners, American English sentences. Reaction times for each condition and each contrast.

Contrast	Condition	RT (ms.)	SD	
Place	Viable	**2038**	761	←
Place	Unviable	1889	768	F[1,25]=18, p<.0001
Place	No-change	**1799**	671	←
Voice	Viable	1958	770	
Voice	Unviable	1887	797	
Voice	No-change	1924	722	

Mean reaction times by subjects are comprised between 1285 ms and 2485 ms (mean RT for n=26: 1920 ms). Analyses of reaction times and detection values did not reveal any interaction of *RT* with the factors *condition* and *type*. We computed the compensation index according to formula (1) for each participant and each item (mean index is 20% for voicing and 43% for place), and used it as a dependent variable in an ANOVA first by participants, then by items. We declared *contrast* as a within-subject (respectively between-item) factor (place vs. voicing). We found a significant effect of *contrast* by participants (not by items), confirming that all subjects behave similarly and compensated significantly more for place assimilation than voicing assimilation ($F_1[1,25]=57$, $p<.0001$; $F_2[1,26]=2.7$, $p>.1$). A t-test revealed that compensation for assimilation was not complete in the place condition, since the compensation index was significantly different from 100% ($t_1(25)=14.6$, $p<.0001$; $t_2(12)=7.6$, $p<.0001$). For the voicing contrast, the index differed significantly from zero ($t_1(25)=5.7$, $p<.0001$; $t_2(14) =2.6$, $p<.05$).

In this experiment, variability in items inhibited various significant effects in our analyses. Looking in greater detail at the pattern of this variability, we see that it mainly concerns voicing items. Place items behave homogenously. Voicing items display an asymmetry between voicing and devoicing items (e.g. *tough* vs. *big*). Compensation was higher for devoicing items: this means that detection (compensation) is higher for 'big fountain' *bi[kf]ountain* (34%) than for 'tough demand' *tou[vd]emand* (8%). The difference between indices for voicing vs. devoicing is significant by participants and items ($F_1[1,25]=23.5$, $p<.0001$; $F_2[1,13]=5.6$, $p=.03$). This could reflect compensation for a process of partial phonetic final devoicing applying in American English (Hyman 1975; Keating 1984: 293). Therefore, for Americans, only the *voicing* items are really non-native. When restricting the analysis to those items, the difference between indices for place and voicing (without *devoicing* items) is very significant by subjects and by items ($F_1[1,25]=34.5$, $p<.0001$; $F_2[1,19]=8.8$, $p<.008$).

Pooled analysis with both experiments on detection rates was performed in order to examine whether listeners' behavior is different across languages, and whether the factor *test-language* interacts with differences due to *contrast type* or to *condition*. Mean detection rate was subjected to a ANOVA with participants as random variable. We declare the factor *test-language* (French or English), as well as both crucial factors *condition* and *contrast*. The factor *test-language* yields no significant main effect, because the directions of effects cancel each other out ($p>.7$). *Test-language* interacts strongly with *contrast* ($F_1[1,42]=54.4$, $p<.0001$) and in a triple interaction also with *condition* ($F_1[2,84]=91.4$, $p<.0001$). This means that both experiments show an opposite pattern of detection, where the test-language strongly influences detection according to contrast type as well as condition.

4.4. Discussion

The main result from Experiment 2 is that American participants listening to American English sentences showed a pattern of results symmetrical to the one observed for French participants listening to French. This result clearly supports the hypothesis that compensation procedures are partly governed by language-specific phonological knowledge. More precisely, we observed that American listeners compensated significantly for changes that correspond to the application of the place assimilation rule in American English. They also compensated for voicing, a process which is not native. However, further analysis of compensation differences between voicing and devoicing revealed that it might be necessary to consider "devoicing" as a native process rather than a non-native one, as opposed to "voicing", which can definitely be considered as non-native, and for which compensation is considerably reduced. In sum, the difference observed in compensation patterns between place and voicing provides further support for the assumption that compensation is driven by language-specific knowledge of phonological processes.

There is one important difference, though, between the French and the American experiments: the amount of compensation for the native rule was larger in French than in American English (65% vs. 46%). This could be due to the fact that place assimilation is less systematic in English than voice assimilation is in French (see Otake, Yoneyama, Culter and van der Lugt 1996, for a similar observation). In other words, the word recognition system for English listeners would be less used to cope with complete place assimilation, than it is used to with complete voicing assimilation in French. When a word is heard in a sentence context, compensation mechanisms are at work, and if they are presented with "optimal" stimuli for which they have been tuned for in the course of language acquisition, they are predicted to be most successful. In our case, the reality of English place assimilation makes our stimuli (because they present rather categorical changes) not optimal for the system to compensate for. This might be slightly different for French stimuli, if the categorical changes we present parallel more closely the reality of French voicing assimilation the system is used to. One could argue that the difference in compensation rate between English and French could originate in the degree of variability in phonetic cues in our stimuli, being more variable in English than in French. Even if this might indeed be present in the stimuli, as indicated by the difference in error rates in the categorization experiment (see below), it does not explain the different compensation patterns in Experiment 1 and 2, for two reasons. First, in case compensation would be the mere reflection of tolerance to cue-uncertainty, one would expect more tolerance in the English case, where cues seem to be more variable, more ambiguous than in French. The difference, however, goes in the opposite direction. Second, one would not expect to find any difference due to condition between viable and unviable condition, i.e. the correct rejection in unviable context (context effects for the native process). For both experiments, the percentage of false alarms in this condition is similar and rather low: for French listeners, voicing yields 06% false alarms, for English listeners, place yields 11%, false alarms in the unviable context. The difference to the respective detection rates in viable conditions is striking (French 65%, English 46%).

The difference observed in the categorization results between English and French – where English listeners make more errors (around 20%) – could reflect a general tendency of phonetic cues to being more variable or less robust in English than in French, especially in this context (see discussion of Experiment 1). Numerous studies have shown systematic differences in the phonetic implementation of particular contrasts between French and English or other languages, with particular attention to the voicing distinction markers (Mack 1982; Kohler 1981, among others). To our knowledge, no study so far examined such systematic differences in cue variability or robustness between English and French, in word-final position before obstruents. Some indirect evidence is found in cross-linguistic studies of intelligibility in time-compressed speech. For a similar compression rate of 50% in English and French sentences, English listeners are able to recall only 44% of the syllables, whereas French listeners listening to compressed French show recall-scores averaging 85% (Mehler et al. 1993; Sebastian-Gallés et al. 2000). In sum, there is a difference in the overall clarity of cues due to particularities of American English and the respective implementation of cues in the particular contexts used. But this cue-robustness difference does not explain the pattern of compensation found in Experiments 1 and 2.

5. General Discussion

The main goal of this study was to investigate the existence of a language-specific phonological knowledge involved in compensation for phonological assimilation. We conducted two experiments, testing two different phonological processes on different languages. Experiment 1 investigated compensation in French native speakers on French stimuli: participants showed more compensation for the voicing contrast than for the place contrast, but only in viable contexts for French voicing assimilation. In

Experiment 2, speakers of American English were tested on American English sentences using the same task: participants compensated more for the place contrast than for the voicing contrast, and only in viable contexts for English place assimilation, thereby presenting symmetrical results from Experiment 1. All these results are supported by additional control experiments, carried out to eliminate the possibility that results could be due to unintentional bias in the stimuli. Excised targets were presented in a forced choice task to new listeners of each language. Words in both change conditions for place and voicing equally were perceived as being different from the form of the target in isolation, meaning that changes were perceived clearly.

Therefore, higher detection rates visible in viable change conditions for the respective native processes is attributable to phonological compensation for assimilation, involving a language-specific knowledge of the processes at work in the language, rather than the language independent use of phonetic cues. Additional support for this view is given by the results presented in Darcy, Peperkamp and Dupoux (2007): In these experiments, listeners – who were also L2 learners of the other language – were presented to both languages, French and American English. French listeners who were beginning learners of English showed the same behavior on both languages, compensating more for voicing assimilation than for place assimilation (69% vs. 40% in French, 64% vs. 37% in English, difference between voicing and place significant). Similarly, American English listeners, who were beginning learners of French (the same participants as in this Experiment 2), showed upon hearing French sentences the same pattern of compensation as they show here, hearing American English sentences (voicing vs. place: 32% vs. 49% in French, and 33% vs. 46% in American English). The fact that they do show a different pattern of compensation on the same stimuli as did the respective native speakers of that language is to be interpreted in the way that these learners still did not acquire the compensation mechanism for that specific process in L2. It excludes the possibility that the observed difference is the result of unintended bias in the stimuli, as here the manipulated variable is only the listener's L1s.

These results converge in showing that compensation is not driven by the unintended acoustic differences between both languages, but rather by the phonological knowledge of the way assimilation works in one language.

Because lexical compensation mechanisms are not sensitive to phonological context, such mechanisms alone cannot explain our results. Similarly, phonetic compensation mechanisms do not rely on familiarity with specific phonological processes, and therefore cannot explain our results either. Nevertheless, we do not think that such mechanisms must necessarily be ruled out. In fact, our data are compatible with the existence of such mechanisms alongside a phonological language-specific, context-sensitive mechanism. The three types of mechanisms would operate at distinct levels of representation, and would all influence subjects' responses in a given task.

To elaborate on our proposal, we postulate that beyond basic auditory processing, speech is initially represented in a universal phonetic format; at this level, language independent mechanisms such as feature parsing may operate (Gow 2001, 2002a; Gow and Im 2004; Gow and Zoll 2002). At the next stage of processing, speech is encoded in a language-specific phonological format; at that level, language-specific mechanisms such as phonological inference to compensate for phonological alternations may operate (our data, Gaskell and Marslen-Wilson 1996, 1998). Finally, such phonological representations are matched against lexical representations for word recognition, in the manner described by multiple activation models (Marslen-Wilson 1987; Marslen-Wilson and Welsh 1978; McClelland and Elman 1986; Norris 1994). Behavioral responses can be influenced by any of these processing levels (as predicted by a multiple readout model). Which level has the greatest influence on behavioral responses depends on many factors, including the task (word identification vs. discrimination), and the nature of the stimuli: whole sentences vs. isolated words or syllables; words vs. nonwords; with large acoustic variations (e.g. across different speakers) or not.

Postulating multiple and cascading compensation mechanisms makes it possible to reinterpret apparently conflicting results from the literature. In the present experiments, we have maximized our chances of observing effects reflecting phonological processing by using words embedded in sentences, and identification across different speakers. Other studies that have used discrimination of nonwords produced by the same speaker have obviously maximized the influence of the phonetic processing level, thereby explaining their finding of universal patterns of compensation.

Gow (2002b) and Gow and Im (2004) reported language independent low-level effects of compensation for voicing assimilation in Hungarian, whether the subjects were native speakers or not (e.g., Korean listeners). These results seem in contradiction with ours. However, it should be noted that these studies used different stimuli from ours: Rather than presenting complete assimilations, they presented ambiguous (multiply articulated) segments, thereby favoring feature parsing. Furthermore, we would like to argue that detecting a word within a sentence across voice changes, the method we used, should force listeners to recode the stimuli at the phonological level and give greater weight to that level in the decision process, as fine acoustic/phonetic details are irrelevant and even interfere with this task. On the other hand, detecting phonemes within bi-syllables without much acoustic variation (their task) may well be more easily performed by paying attention to the phonetic level of representation. According to this interpretation, both our results and those of Gow (2002b) and Gow and Im (2004) can be explained by the same multiple readout model; simply, their experiments induce responses predominantly based on phonetic representations and therefore reflect universal phonetic processes, whereas our experiments (and those of Gaskell and Marslen-Wilson 1996, 1998) induce responses based primarily on phonological representations, therefore reflecting language-specific abstract phonological processes.

Restated within this framework, our results show that the phonological level is responsible for most of the effects observed in our experiments, as it is the only level where both context-sensitive and language-specific effects may arise. But even before this phonological inference mechanism applies, some degree of universal feature parsing may occur, prompted by e.g. homorganic clusters. This effect could explain the small, but non-null compensation for voicing assimilation by English listeners, and for place assimilation by French listeners. Finally, lexical compensation mechanisms may also have played a role in our experiments. Such a mechanism would generate a global tendency to detect the target based on phonological proximity. It could be responsible in part for the error rate in the unviable context (across the experiments from 6% to 18%).

Although our results make clear that a context-sensitive phonological knowledge of processes is at work, they leave open the question of whether such a mechanism operates at a strictly sub-lexical level (i.e., before lexical access) or whether it is implemented as a more sophisticated, context-sensitive version of a lexical compensation mechanism. Further research involving nonwords will be needed to answer that question.

Author's Note

We thank Sharon Peperkamp and Stig Eliasson for useful discussion and feedback, Jim Morgan, Joe Pater and John Kingston for providing testing facilities. As well, we are grateful to the volume editors Caroline Féry, Ruben van de Vijver and Frank Kügler, for providing helpful comments and discussion. This work was supported by a Ph.D. fellowship of the Landesgraduiertenförderung Rheinland Pfalz from Mainz University (I.D.), a Franco- German mobility fellowship from the Robert Bosch foundation (I.D.), a Marie Curie fellowship of the European Community programme Quality of Life QLG1-CT-1999-51305 (F.R.), an APN grant ”Les surdités phonologiques: Etudes interlangues”, 2000 (E.D. / Sharon Peperkamp), and a grant from CNRS “Acquisition phonologique précoce : algorithmes et simulations”, 2002 (E.D. / Sharon Peperkamp).

Appendix

French words used in experiment 1

American words used in experiment 2

Notes

1. Here, [nw] means that the word underwent an assimilatory change, and became a nonword.
2. This constraint made it necessary to include geminate clusters in the place set, otherwise the place agreement would have also produced violation of the voicing agreement constraint in French. In order to balance both sets, we also included the same number of geminates in the voicing set. The speaker produced all geminates as a single long consonant, without release in between. The same constraint has been obeyed for English stimuli sets.
3. For this and the following experiments, all speakers were trained until they are familiar with the nonwords, and able to pronounce all sentences in a natural way. We avoided cross splicing due to the difficulties to match whole sentences with respect to prosody and speech rate.
4. Reaction times were collected for a “yes response”. Restriction to these two conditions is due to the fact that only those conditions present sufficient response rates in order to allow for a valid estimation of reaction times

Table 7: French words used in experiment 1 (Place).

Target	Gloss	Un-changed form	Changed form	No-change Context	Unviable Context	Viable Context
				Place		
bête	(beast)	[bɛt]	[bɛp]	nuisible ‘cumbrous’ [nɥizibl]	feroce ‘ferocious’ [feʁɔs]	poilue ‘hairy’ [pwalyː]
boîte	(box)	[bwat]	[bwak]	marron ‘brown’ [maʁɔ̃]	fermee ‘closed’ [fɛʁmeː]	carrée ‘square’ [kaʁeː]
botte	(boot)	[bɔt]	[bɔp]	montantes ‘high’ [mɔ̃tɑ̃t]	rayées ‘striped’ [ʁɛjeː]	pointue ‘spiky’ [pwɛ̃tyː]
chouette	(owl)	[ʃwɛt]	[ʃwɛk]	malade ‘sick’ [malad]	sauvage ‘wild’ [sovaʒ]	craintive ‘frightened’ [kʁɛ̃tiv]
dune	(dune)	[dyn]	[dym]	lointaine ‘remote’ [lwɛ̃tɛn]	sauvage ‘wild’ [sovaʒ]	brumeuse ‘brumous’ [bʁymøz]
guide	(guide)	[gid]	[gib]	raciste ‘racist’ [ʁasist]	vulgaire ‘vulgar’ [vylgɛʁ]	bourru ‘grouchy’ [buʁy]
lune	(moon)	[lyn]	[lym]	jaune ‘yellow’ [ʒon]	rousse ‘red’ [ʁus]	pâle ‘pale’ [pɑl]
mode	(fashion)	[mɔd]	[mɔg]	locale ‘local’ [lokal]	zoulou ‘Zulu’ [zulu]	guerrière ‘combat’ [gɛʁjɛʁ]
moine	(monk)	[mwan]	[mwam]	rusé ‘wily’ [ʁyze]	serviable ‘helpful’ [sɛʁvjabl]	bavard ‘talkative’ [bavaʁ]
prune	(plum)	[pʁyn]	[pʁym]	juteuses ‘juicy’ [ʒytøz]	sucrées ‘sweet’ [sykʁeː]	pourries ‘rotten’ [puʁiː]
reine	(queen)	[ʀɛn]	[ʀɛm]	généreuse ‘generous’ [ʒeneʁøz]	respectée ‘respected’ [ʁɛspɛkteː]	paresseuse ‘lazy’ [paʁɛsøz]
ride	(wrinkle)	[ʀid]	[ʀig]	légère ‘light’ [leʒɛʁ]	discrète ‘discreet’ [diskʁɛt]	gracieuse ‘graceful’ [gʁasjøz]
ruine	(ruin)	[ʀɥin]	[ʀɥim]	romaine ‘Latin’ [ʁomɛn]	célèbre ‘famous’ [selɛbʁ]	baroque ‘baroque’ [baʁɔk]
stade	(stadium)	[stad]	[stab]	rénové ‘renovated’ [ʁenove]	démodé ‘outdated’ [demode]	bétonné ‘concrete’ [betone]
trone	(throne)	[tʁon]	[tʁom]	rocheux ‘rocky’ [ʁoʃø]	royal ‘royal’ [ʁwajal]	princier ‘princely’ [pʁɛ̃sje]
zone	(zone)	[zon]	[zom]	rurale ‘rural’ [ʁyʁal]	fluviale ‘riverine’ [flyvjal]	portuaire ‘harbor’ [pɔʁtɥɛʁ]

Table 8: French words used in experiment 1 (Voicing).

Target	Gloss	Un-changed form	Changed form	No-change Context	Unviable Context	Viable Context
				Voicing		
badge	(badge)	[badʒ]	[batʃ]	métallique ‘metallic’ [metalik]	ravissant ‘charming’ [ʁavisɑ̃]	parfumé ‘perfumed’ [paʁfyme]
cape	(cape)	[kap]	[kab]	longue ‘long’ [lɔ̃g]	neuve ‘new’ [nøv]	grise ‘grey’ [gʁiz]
chèque	(check)	[ʃɛk]	[ʃɛg]	mensuel ‘monthly’ [mɑ̃sɥɛl]	reçu ‘received’ [ʁəsy]	volé ‘stolen’ [vole]
couche	(layer)	[kuʃ]	[kuʒ]	neigeuse ‘snow’ [neʒøz]	marron ‘brown’ [maʁɔ̃]	jaunie ‘yellowed’ [ʒoniː]
coude	(elbow)	[kud]	[kut]	meurtri ‘injured’ [mœʁtʁi]	raidi ‘rigid’ [ʁɛdi]	tordu ‘twisted’ [tɔʁdy]
cuve	(tank)	[kyv]	[kyf]	mobile ‘mobile’ [mobil]	remplie ‘full’ [ʁɑ̃pliː]	fendue ‘ripped’ [fɑ̃dyː]
faute	(error)	[fot]	[fod]	majeure ‘major’ [maʒœʁ]	légère ‘light’ [leʒɛʁ]	discrète ‘discreet’ [diskʁɛt]
globe	(globe)	[glɔb]	[glɔp]	miroitant ‘mirroring’ [miʁwatɑ̃]	lumineux ‘luminous’ [lyminø]	pailleté ‘sequined’ [pajəte]
lac	(lake)	[lak]	[lag]	limpide ‘clear’ [lɛ̃pid]	nordique ‘Nordic’ [nɔʁdik]	gelé ‘frosted’ [ʒəle]
lave	(lava)	[lav]	[laf]	mouvante ‘moving’ [muvɑ̃t]	rugueuse ‘cragged’ [ʁygøz]	pateuse ‘pasty’ [pɑtøz]
nappe	(tablecloth)	[nap]	[nab]	rayée ‘striped’ [ʁɛjeː]	rustique ‘rustic’ [ʁystik]	brodée ‘embroidered’ [bʁodeː]
neige	(snow)	[nɛʒ]	[nɛʃ]	mouillée ‘wet’ [mujeː]	marron ‘brown’ [maʁɔ̃]	poudreuse ‘powder’ [pʊdʁøz]
nuage	(cloud)	[nɥaʒ]	[nɥaʃ]	rosés ‘rosy’ [ʁoze]	nacrés ‘pearly’ [nakʁe]	chargés ‘loaded’ [ʃaʁʒe]
plaque	(plate)	[plak]	[plag]	noircie ‘blacked’ [nwarsiː]	rouillée ‘rust’ [ʁujeː]	brillante ‘shiny’ [bʁijɑ̃t]
robe	(dress)	[ʀɔb]	[ʀɔp]	rouge ‘red’ [ʁuʒ]	noire ‘black’ [nwaʁ]	sale ‘dirty’ [sal]
route	(road)	[ʀut]	[ʀud]	magnifique ‘beautiful’ [maɲifik]	nationale ‘main’ [nasjonal]	dangereuse ‘dangerous’ [dɑ̃ʒəʁøz]

Table 9: American words used in experiment 2 (Place).

Target	Unchanged form	Changed form	No-change Context	Unviable Context	Viable Context
			Place		
bad	[bæd]	[bæb]	[dɪʃ] dish	[lʌnʃ] lunch	[bɪə] beer
fat	[fæt]	[fæp]	[mʌŋkiː] monkey	[skwɪɹɪl] squirrel	[pʌpi] puppy
great	[gɹeɪt]	[gɹeɪk]	[faɪt] fight	[mætʃ] match	[kruːz] cruise
mad	[mæd]	[mæb]	[mɒðɜː] mother	[dɒtɜː] daugther	[bɹɑðɜː] brother
red	[ɹed]	[ɹeg]	[nekleɪs] necklace	[lɪpstɪk] lipstick	[glæsɪz] glasses
sad	[sæd]	[sæb]	[mʊviː] movie	[nɒvəl] novel	[bəleɪ] ballet
sweet	[swiːt]	[swiːk]	[ʃʌklɛtt] chocolate	[lɪkjɜː] liqueur	[kʌkteɪl] cocktail
wet	[wet]	[wep]	[ʃuːz] shoes	[sɒks] socks	[pænts] pants
clean	[kliːn]	[kliːm]	[fɔːk] fork	[spuːn] spoon	[pæn] pan
fun	[fʌn]	[fʌŋ]	[deɪ] day	[naɪt] night	[geɪm] game
green	[gɹiːn]	[gɹiːŋ]	[veɪz] vase	[tʃæː] chair	[kʌp] cup
lean	[liːn]	[liːm]	[laɪn] line	[ʃeɪp] shape	[bæk] back
own	[əʊn]	[əʊm]	[laɪf] life	[tʃɔɪs] choice	[plæn] plan
plain	[pleɪn]	[pleɪŋ]	[tʃæpəls] chapels	[tʃɜːtʃɪz] churches	[kɒndəʊz] condos
tan	[tæn]	[tæm]	[skɑːf] scarf	[ʃɜːt] shirt	[belt] belt
thin	[θɪn]	[θɪm]	[nəʊtbʊk] notebook	[liːflet] leaflet	[pækɪt] packet

Table 10: American words used in experiment 2 (Voicing).

Target	Unchanged form	Changed form	No-change Context	Unviable Context	Viable Context
			Voicing		
big	[bɪg]	[bɪk]	[laɪthaʊs] lighthouse	[ɹɪvɜː] river	[faʊntɪn] fountain
blind	[blaɪnd]	[blaɪnt]	[leɪdi] lady	[lɔːjɜː] lawyer	[tʃelɪst] cellist
brave	[bɹeɪv]	[bɹeɪf]	[məɹiːn] marine	[laɪfgɑːd] lifeguard	[faɪɜːmæn] fireman
drab	[dɹæb]	[dɹæp]	[laɪtɪŋ] lighting	[meɪkʌp] make-up	[peɪntɪŋ] painting
good	[gʊd]	[gʊt]	[lʊks] looks	[lʌk] luck	[fɹendz] friends
huge	[hjuːdʒ]	[hjuːtʃ]	[manɜː] manor	[mænʃən] mansion	[fɒɹest] forest
mild	[maɪld]	[maɪlt]	[naɪts] nights	[ɹeɪn] rain	[spɹɪŋ] spring
wise	[waɪz]	[waɪs]	[liːdɜː] leader	[ɹaɪtɜː] writer	[tiːtʃɜː] teacher
best	[best]	[bezd]	[muːv] move	[ɹʌn] run	[deɪ] day
black	[blæk]	[blæg]	[ɹʌg] rug	[ɹæg] rag	[glʌv] glove
cheap	[tʃiːp]	[tʃiːb]	[lʌnʃ] lunch	[ɹuːm] room	[dɹɪŋk] drink
flat	[flæt]	[flæd]	[ɹɑːft] raft	[ɹʌk] rock	[dæm] dam
French	[fɹentʃ]	[fɹendʒ]	[meɪd] maid	[nɜːs] nurse	[gaɪd] guide
nice	[naɪs]	[naɪz]	[mædəʊz] meadows	[ɹeɪlɪŋz] railings	[gɑːdənz] gardens
thick	[θɪk]	[θɪg]	[ɹəʊp] rope	[laɪn] line	[bɑː] bar
tough	[tʌf]	[tʌv]	[lesən] lesson	[ɹɪkwest] request	[dɪmænd] demand

References

Archangeli, Diana
1988 Aspects of underspecification theory. *Phonology* 5: 183-207.

Bailey, P. J. and Quentin Summerfield
1980 Information in speech: Some observations on the perception of [s]-stop clusters. *Journal of Experimental Psychology: Human Perception and Performance* 6: 536-563.

Barry, Martin
1992 Palatalisation, assimilation and gestural weakening in connected speech. *Speech Communication* 11: 393-400.

Best, Catherine T., B. Morrongiello and R. Robson
1981 Perceptual equivalence of acoustic cues in speech and nonspeech perception. *Perception and Psychophysics* 29: 191-211.

Casagrande, Jean
1984 *The sound system of French.* Washington D. C.: Georgetown University Press.

Church, Kenneth
1987 Phonological parsing and lexical retrieval. *Cognition* 25: 53-69.

Coenen, Else, Pienie Zwitserlood and Jens Bölte
2001 Variation and assimilation in German: consequences for lexical access and representation. *Language and Cognitive Processes* 16: 535-564.

Content, Alain, P. Mousty and Monique Radeau
1990 BRULEX. Une base de données lexicales informatisée pour le français écrit et parlé. *L'Année Psychologique* 90: 551-566.

Cutler, Anne and Dennis Norris
1979 Monitoring sentence comprehension. In *Sentence Processing: Psycholinguistic Studies presented to Merrill Garrett;* W. E. Cooper and E. C. T. Walker (eds.), 113-134. Hillsdale N.J.: Erlbaum.

Darcy, Isabelle, Sharon Peperkamp and Emmanuel Dupoux
2007 Bilinguals play by the rules. Perceptual compensation for assimilation in late L2-learners. In *Laboratory Phonology 9*, Jennifer Cole and José I. Hualde (eds.), 411-442. Berlin: Mouton de Gruyter.

Dell, François
1995 Consonant clusters and phonological syllables in French. *Lingua* 95: 5-26.

Ellis, Lucy and William J. Hardcastle
2002 Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data. *Journal of Phonetics* 30: 373-396.

Féry, Caroline
2003 Gradedness in the segmental correlates of focus and prosodic phrasing

in French. In *Nouveaux départs en phonologie*, Trudel Meisenburg and Maria Selig (eds.) 161-181. Tübingen: Narr.

Fowler, Carol A.
1996 Listeners do hear sounds, not tongues. *Journal of the Acoustical Society of America* 99: 1730-1741.

Fowler, Carol A. and J. M. Brown
2000 Perceptual parsing of acoustic consequences of velum lowering from information for vowels. *Perception and Psychophysics* 62: 21-32.

Gaskell, Gareth
2003 Modelling regressive and progressive effects of assimilation in speech perception. *Journal of Phonetics* 31: 447-463.

Gaskell, Gareth, Mary Hare and William D. Marslen-Wilson
1995 A connectionist model of phonological representation in speech perception. *Cognitive Science* 19: 407-439.

Gaskell, Gareth and Marslen-Wilson, William D.
1996 Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance* 22: 144-158.

Gaskell, Gareth and Marslen-Wilson, William D.
1998 Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance* 24: 380-396.

Gaskell, Gareth and Marslen-Wilson, William D.
2001 Lexical ambiguity resolution and spoken word recognition: bridging the gap. *Journal of Memory and Language* 44: 325-349.

Gow, David W.
2001 Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language* 45: 133-159.

Gow, David W.
2002a Does English coronal place assimilation create lexical ambiguity? *Journal of Experimental Psychology: Human Perception and Performance* 28: 163-179.

Gow, David W.
2002b Integrating feature cues over time to resolve ambiguity. In *Temporal Integration in the perception of speech TIPS*, Sarah Hawkins and Noël Nguyen (eds.), P1-10. ISCA Tutorial and Research Workshop ITRW, Aix-en-Provence, France, April 8-10, 2002. http://www.isca-speech.org/archive/tips_02

Gow, David W.
2003 Feature Parsing: Feature cue mapping in spoken word recognition. *Perception and Psychophysics* 65: 575-590.

Gow, David W. and P. C. Gordon
1995 Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception and Performance* 21: 344-359.

Gow, David W. and A. M. Im
2004 A cross-linguistic examination of assimilation context effects. *Journal of Memory and Language* 51: 279-296.

Gow, David W. and Cheryl Zoll
2002 The Role of Feature Parsing in Speech Processing and Phonology. In *Phonological Answers and their corresponding questions*, Anikó Csirmaz, Zhiqiang Li, Andrew Nevins, Olga Vaysman and Michael Wagner (eds.), 55-68. MIT Working Papers in Linguistics 42.

Grainger, J. and A. M. Jacobs
1996 Orthographic processing in visual word recognition: a multiple readout model. *Psychological review* 103: 518-565.

Gumnior, Heidi, Pienie Zwitserlood and Jens Bölte
2005 Assimilation in existing and novel German compounds. *Language and Cognitive Processes* 20: 465-488.

Hallé, Pierre, Céline Chereau and Juán Segui
2000 Where is the /b/ in “absurde” [apsyrd]? It is in French listeners’ minds. *Journal of Memory and Language* 43: 618-639.

Hayes, Bruce
1984 The phonetics and phonology of Russian voicing assimilation. In *Language sound structure*, Mark Aronoff and Richard T. Oehrle (eds.), 318-328. Cambridge MA: MIT Press.

Hodgson, P. and Joanne L. Miller
1996 Internal structure of phonetic categories: Evidence for within-category trading relations. *Journal of the Acoustical Society of America* 100: 565-576.

Hyman, L. M.
1975 *Phonology: Theory and analysis.* New York: Holt, Rinehart and Winston.

Jakobson, Roman
1956 Die Verteilung der stimmhaften und stimmlosen Geräuschlaute im Russischen. In *Festschrift für Max Vasmer zum 70. Geburtstag*, Margarete Woltner and Herbert Bräuer (eds.), 199-202. Wiesbaden: Harassowitz.

Jun, Jongho
2004 Place assimilation. In *Phonetically Based Phonology*, Bruce Hayes, Robert Kirchner and Donca Steriade (eds.), 58-86. Cambridge: Cambridge University Press.

Keating, P. A.
1984 Phonetic and phonological representation of stop consonant voicing. *Language* 60: 286 - 319.

Kingston, John and Randy Diehl
1994 Phonetic knowledge. *Language* 70: 419-454.

Kiparsky, Paul
1985 Some consequences of lexical phonology. *Phonology Yearbook* 2: 85-138.

Koster, C. J.
1987 *Word recognition in foreign and native language*. Dordrecht, The Netherlands: Foris.

Kohler, Klaus
1981 Contrastive Phonology and the acquisition of phonetic skills. *Phonetica* 38: 213-226.

Kohler, Klaus
1984 Phonetic features in phonology: the feature fortis-lenis. *Phonetica* 41: 150-174.

Lahiri, Aditi and William D. Marslen-Wilson
1991 The mental representation of lexical form: a phonological approach to the recognition lexicon. *Cognition* 38: 245-294.

Lahiri, Aditi and William D. Marslen-Wilson
1992 Lexical processing and phonological representation. In *Papers in Laboratory Phonology 2*, D. R. Ladd and Gerry J. Docherty (eds.), 229-260. Cambridge: Cambridge University Press.

Lahiri, Aditi and Henning Reetz
2002 Underspecified recognition. In *Papers in Laboratory Phonology 7*, Carlos Gussenhoven and Natasha Warner (eds.), 637-675. Berlin: Mouton DeGruyter.

Liberman, Alwin M., F. S. Cooper, D. P. Shankweiler and M. Studdert-Kennedy
1967 Perception of the speech code. *Psychological Review* 74: 431-461.

Lotto, Andrew J., Keith R. Kluender and Lori L. Holt
1997 Perceptual compensation for coarticulation by Japanese quail Coturnix coturnix japonica. *Journal of the Acoustical Society of America* 102: 1134-1140.

Mack, Molly
1982 Voicing-dependent vowel duration in English and French: Monolingual and bilingual production. *The Journal of the Acoustical Society of America* 71: 173-178.

Marslen-Wilson, William D.
1987 Functional parallelism in spoken word-recognition. *Cognition* 25: 71-102.

Marslen-Wilson, William D., Andy Nix and Gareth Gaskell
1995 Phonological variation in lexical access: abstractness, inference and English place assimilation. *Language and Cognitive Processes* 10: 285-308.

Marslen-Wilson, William D. and A. Welsh
1978 Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology* 10: 29-63.

McClelland, J. L. and J. L. Elman
1986 The TRACE model of speech perception. *Cognitive Psychology* 18: 1-86.

Mehler, Jacques, Nuria Sebastian, Gerry Altmann, Emmanuel Dupoux, Anne Christophe and Christophe Pallier
1993 Understanding compressed sentences: the role of rhythm and meaning. *Annals of the New York Academy of Sciences* 682: 272-282.

Mitterer, Holger and Leo Blomert
2003 Coping with phonological assimilation in speech perception: Evidence for early compensation. *Perception and Psychophysics* 65: 956 - 969.

Mitterer, Holger, Valeria Csépe and Leo Blomert
2003 Compensation for phonological assimilation in perception: Evidence from Hungarian liquid assimilation. In *Proceedings of the15th International Congress of Phonetic Sciences*, M. J. Solé, D. Recasens and J. Romero (eds.), 2321-2324. Barcelona, 3-9 August 2003.

Nolan, Francis
1992 The descriptive role of segments: evidence from assimilation. In *Papers in Laboratory Phonology 2*, D. R. Ladd and Gerry J. Docherty (eds.), 261-280. Cambridge: Cambridge University Press.

Norris, Dennis
1994 Shortlist: a connectionist model of continuous speech recognition. *Cognition* 52: 189-234.

Otake, T., K. Yoneyama, Anne Cutler and Arie van der Lugt
1996 The representation of Japanese moraic nasals. *Journal of the Acoustical Society of America* 100: 3831-3842.

Pallier, Christophe, Emmanuel Dupoux and Xavier Jeannin
1997 EXPE: an expandable programming language for on-line psychological experiments. *Behavior Research Methods, Instruments and Computers* 29: 322-327. http://www.lscp.net/expe.

Parker, E. M., Randy C. Diehl and Keith R. Kluender
1986 Trading relations in speech and nonspeech. *Perception and Psychophysics* 39: 129-142.

Pulleyblank, Douglas
1988 Underspecification, the feature hierarchy and Tiv vowels. *Phonology* 5: 299-326.

Quené, Hugo, M. van Rossum and M. van Wijck
1998 Assimilation and anticipation in word perception. In *Proceedings of the Fifth International Conference on Spoken Language Processing*, volume 3: 871-874. Sydney, December 1998.

Repp, Bruno H.
1982 Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin* 92: 81-110.

Rigault, André
1967 L'assimilation consonantique de sonorité en français: étude acoustique et perceptuelle. In *Proceedings of the 6th International Congress of Phonetic Sciences*, B. Hála, M. Romportl and P. Janotà (eds.), 763-766. Prague: Academia.

Roca, Iggy and Johnson, W.
1999 *A Course in Phonology*. Oxford: Blackwell.

Samuel, Arthur G.
1981 Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General* 110: 474-494.

Samuel, Arthur G.
1996 Does lexical information influence the perceptual restoration of phonemes? *Journal of Experimental Psychology: General* 125: 28-51.

Samuel, Arthur G.
2001 Knowing a word affects the fundamental perception of the sounds within it. *Psychological Science* 12: 348-351.

Sebastian-Gallés, Nuria, Emmanuel Dupoux, Albert Costa and Jacques Mehler
2000 Adaptation to time-compressed speech: phonological determinants. *Perception and Psychophysics* 62: 834-842.

Sinnott, J. M. and T. A. Saporita
2000 Differences in American English, Spanish and monkey perception of the say-stay trading relation. *Perception and Psychophysics* 62: 1312-1319.

Snoeren, Natalie, Pierre Hallé and Juán Segui
2006 A voice for the voiceless: Production and perception of assimilated stops in French. *Journal of Phonetics* 34: 241-268.

Stevens, Kenneth
1998 *Acoustic phonetics*. Cambridge MA: MIT Press.

Summerfield, Quentin and M. Haggard
1977 On the dissociation of spectral and temporal cues to the voicing distinction in initial stop consonants. *Journal of the Acoustical Society of America* 62: 435-448.

Treiman, M.
1999 There are two types of psychometric functions: A theory of cue combination in the processing of complex stimuli with implications for categorical perception. *Journal of Experimental Psychology: Human Perception and Performance* 25, 517-546.

Trubetzkoy, Nikolai S.
1958 *Grundzüge der Phonologie*. Göttingen: Vandenhoeck and Ruprecht [first edition 1939].

Weber, Andrea
2001 Help or hindrance: How violation of different assimilation rules affects spoken language processing. *Language and Speech* 44: 95-118.

Weber, Andrea
2002 Assimilation violation and spoken-language processing: A supplementary report. *Language and Speech* 45: 37-46.

Wells, J. C.
1982 *Accents of English I*. Cambridge: Cambridge University Press.

Wetzels, W. Leo and Joan Mascaró
2001 The typology of voicing and devoicing. *Language* 77: 207-244.

Wiese, Richard
1996 *The phonology of German*. Oxford: Clarendon Press.

Wilson, M. D.
1988 MRC Psycholinguistic Database: Machine-usable dictionary, version 2.0. *Behavior Research Methods, Instrumentation and Computers* 20: 6-10.

Wright, Richard
2004 A review of perceptual cues and cue robustness. In *Phonetically Based Phonology*, Bruce Hayes, Robert Kirchner and Donca Steriade (eds.), 34-57. Cambridge: Cambridge University Press.

MRC Psycholinguistic Database :
http://www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm retrieved 11.07.2006, 13h00

The erosion of a variable process.
The case of n-deletion in Ripuarian and Limburg dialects of Dutch

Frans Hinskens

Abstract

This paper addresses the phenomenon of variable final n-deletion after lax vowels in monosyllabic words as such and its usage in in-group conversations. The occurrence of this deletion process is confined to a group of Riparian and Limburg dialects of Dutch. The phenomenon (its linguistic *raison d'être* as well as its variable use) and the apparent time erosion of the phenomenon are explained from internal and social perspectives. The phenomenon is complex, as it is simulaneously conditioned by a multitude of internal factors.

An attempt is made to locate the internal (phonological and prosodic) and social forces that variable n-deletion is subject to. The combined effects of the interaction between internal and social forces on the one hand and the permanent contact with non-n-deleting varieties on the other makes the phenomenon vulnerable and explains its gradually (and, as yet, conditionally) falling into disuse.

The outcomes of quantitative analyses of the variability in n-deletion enrich the ('Labovian') model for the quantitative study of language variation and change by adding insights into processes leading to the *demise* of linguistic variation and cross-dialectal convergence to this model, which was mainly designed to study the distribution and diffusion of linguistic variation.

1. Introduction[1]

Although processes of convergence and divergence of traditional dialects are not new phenomena, they seem to have gained momentum in twentieth century Europe (Auer and Hinskens 1996; Auer, Hinskens and Kerswill 2005). Dialect convergence and divergence are, of course, not uniquely European phenomena. In North America there are English dialects which have converged towards the standard variety to such an extent that they are on the verge of being wiped out (see e.g. Wolfram and Schilling Estes 1995). On the other hand, massive dialectal diversification is taking place; Labov 1991 (cf. also Labov, Ash and Boberg 1995; 2006) points out that in North America divergence seems predominant. In Europe, however, the more general case seems to be for related dialects to converge, with processes of dialect levelling playing a crucial role in this type of language change. The overall difference between North America (mostly divergence) and Europe (mainly convergence) might be related to the difference in time-depth and the linguistic and sociolinguistic effects it may have - although this is largely speculation as yet.[2]

In this contribution, a study of ongoing processes of dialect levelling in an originally urban Limburg dialect of Dutch will be introduced. After a few considerations concerning the relationship between dialect levelling and language change (§ 2), a brief sketch of the research area will be given (§ 3).

One of the dialect features studied is the variable deletion of final /n/ in monosyllabic words when preceded by a short vowel (type *in*, ‘in’, *ben*, ‘am’, *dan*, ‘then; than’ etc.). In §§ 4 and 5 a description and a formal account of the deletion and its complicated internal conditioning, as well as of the seemingly complicated blockings of the process will be presented. The model of grammar applied is basically Lexical Phonology enriched, on the subsegmental level, with insights from Feature Geometry and, for the phonology-syntax connection, with Prosodic Phonology. After a brief account of the methods applied for the collection and quantitative analyses of the data (§ 6), the findings regarding linguistic as well as extralinguistic aspects of the levelling out, in casu the gradual loss, of this deletion rule will be discussed (§ 7).

Although on an overall level the n-deletion rule is not endangered, it appears to be gradually but unmistakably falling into disuse in several specific linguistic conditions. Interestingly, application of the rule in these linguistic conditions appears to be in a process of becoming socio-stylistically marked. In other words, n-deletion is probably in the process of changing from a dialect feature in the traditional sense into a sociolect feature - or, to put it differently, from an indicator into a marker (Labov 1972).

In the discussion of these findings (§ 8), attention will be paid to some implications of insights into processes of dialect levelling and convergence for models of language change as well as to the relationship between sociolinguistics and formal linguistic theory.

2. Dialect levelling and language change

Dialect levelling can be defined as the process of the reduction of formal variation. This succinct definition captures both quantitative, internal variation and variation between varieties of a language - say, dialects. In the latter case, dialect levelling may lead to the formation of a 'koiné'. The levelling out of variation between a dialect and the related standard language on the hand and of variation across related dialects on the other need not be mutually dependent (Hinskens 1992: § 12.2.1; 1998).

In line with the above definition, the conceptual relationships between language variation, dialect levelling and linguistic change can be envisaged as follows. A process of language change that has not come to completion in some respect leaves behind language variation. Internal variation, on the other hand, can lead to linguistic change - or rather, language variation can be the synchronic reflection of an ongoing process of linguistic change. Dialect levelling is the process which reduces language variation - it is hence a special type of linguistic change.

Less than a few centuries ago, in the age when -at least in the Western world- there was little or nothing corresponding to the modern concept of 'standard language', there were only what is nowadays referred to as 'dialects'. What was language change like in those days? Of course there has always been language change resulting from internal pressures, such as phonetically motivated sound change, change resulting from some structural imbalance because of e.g. asymmetries in the vowel system, Sapir's 'drift', the tendency to keep paradigms regular and transparent - e.g. to restore paradigmatic regularity after a sound change had destroyed it etc. Another inexhaustible source of language change is the modular organization of grammar and phonology. These are purely *internal* motivations.

As far as *externally* motivated language change is concerned, the mixing and borrowing resulting from the prolonged contact between dialects have probably always been very prominent phenomena. According to Ferdinand Wrede (see e.g. 1919: 10-13) and Theodor Frings, two of the protagonists in the history of German dialectology, 'Mischung' and 'Ausgleich', i.e. mixing and cross-dialectal levelling, are the key mechanisms that destroy regularity and the alleged exceptionlessness of sound laws.

As convergence processes, mixing and levelling across dialects probably played leading roles in the shaping of the type of koiné that many contemporary standard languages basically are. Historically, the Dutch standard language, for instance, is a koiné of some Flemish and Brabant and many Hollandic dialect features.

In the light of these considerations, the study of processes of dialect levelling in the wider sense can be located at the crossroads of the study of language variation on the one hand and research into the structural consequences of language contact on the other.

3. The research area

The present contribution zooms in on an investigation of the process of the levelling out of a feature of the local dialect of Rimburg. Rimburg is a small village (with less than 1.000 inhabitants) in the southeast of the Dutch province of Limburg . It lies in the heart of an urbanized area. Rimburg is not far from Heerlen and Kerkrade; it lies on the German border, approximately 10 kms from the German city of Aachen (Aix-la-Chapelle).

Rimburg is located immediately east of the Benrath line, the famous isogloss bundle which marks the western border of (the ‘Rhenish Fan’, the transition zone north and west of) the area where the Old High German voiceless stops have been changed into affricates and fricatives in certain positions, the so-called Second Consonant Shift or Old High German Consonant Shift. This isogloss bundle hence divides the High-German dialects, which have undergone this historical change, and the Low-German ones, which have not (Wolf 1983). In dialect-geograpy, the dialects east of the Benrath line are commonly referred to as Ripuarian dialects; west of the line, Limburg dialects are spoken.

In recent history, the southeastern part of the country has been of a ‘natural laboratory’, or maybe even a ‘pressure cooker’, of demographic, social and cultural changes. These changes were brought about by the very rapid industrialization (coalmining) of this area in the first decades of the 20th century. The industrialization created job opportunities, which led to considerable immigration - from foreigners (especially from Eastern Europe) and even more so from people from other parts of the country. The migration led, in turn, to urbanization. All coalmines were closed down between 1966 and 1976, but the effects of the industrialization remain. At present, with respect to the number of inhabitants, the Heerlen- Kerkrade agglomeration ranks among the largest in the Netherlands. Its average density of population is approximately three times the national average.[3]

4. Word-final n-deletion. The process and its internal motivation

One of the features distinguishing the Rimburg dialect (as well as a number of neighbouring dialects in the Heerlen-Kerkrade area)[4] is n-deletion, specifically, the variable deletion of final /n/ in monosyllabic words in which final /n/ is preceded by a short vowel. The process of n-deletion constitutes variation between dialects, i.e between the dialects that have the rule and the ones that do not. At the same time, this process of n-deletion constitutes a case of internal variation, since n-deletion is a variable process.

A list of the words that can constitute input to the rule is given in (1):

(1)

PREPOSITIONS				
dial.	/vɑn/	std.lg.	*van*	‘of’, ‘from’[5]
	/ɪn/		*in*	‘in’
	/ɑn/		*aan*	‘on’ (iff realized with a short vowel)
CONJUNCTION				
	/ɛn/		*en*	‘and’
ADVERBS				
	/dɑn/		*dan*	‘then’, ‘than’
	/nun/		*nu, nou*	‘now’
	/hɪn/		*heen*	‘away’, ‘to’
Vfin (1 st person sing. present indic.)				
	/ɣ˧ɔ̞n/		*ga*	‘go’
	/kRin/		*krijg*	‘get’
	/hɑn/		*heb*	‘have’
	/dɔ̞n/		*doe*	‘do’
	/zɑn/		*zeg*	‘say’
	/zin/		*zie*	‘see’
	/ʃtɔ̞n/		*sta*	‘stand’
	/ʃlɔ̞n/		*sla*	‘beat’
	/kɑn/		*kan*	‘can’
	/kɪ̝n/		*ken*	‘know’
	/bɪ̝n/		*ben*	‘am’
Vfin (3 rd person sing. present indic.)[6]				
	/kɑn/		*kan*	‘can’
Vinfin				
	/hɑːn/		*hebben*	‘to have’

In these items, n-deletion, although quantitatively gradient (or ‘optional’), is phonetically abrupt in that it applies without leaving a trace. In case the /n/ is deleted, no nasalization or compensatory lengthening of the vowel occurs (as a native speaker of one of the relevant dialects, the present author can tell).

The prepositions /vɑn/, /ɪn/ and /ɑn/ function as adverbs under certain grammatical conditions; /ɣ˧ɔ̝n/, /kRin/, /dɔn/, /zin/, /ʃtɔ̝n/ and /ʃlɔ̝n/ are forms of athematic verbs, /zɑn/ is athematic in the pres.indic. inflection.

The list in (1) is exhaustive from the point of view of the input to the rule. It is not exhaustive from the point of view of the lexicon; the dialect has some thirty other words of this phonological make-up which *never* undergo final n-deletion. Examples are:

(2)

NOUNS			
	/kɑn/	*kan*	‘jug’, ‘pitcher’
	/dɛn/	*den*	‘pine tree’
	/pɛn/	*pen*	‘pen’
	/dɪn/	*dorsloer*	‘threshing floor’
	/ʏn/	*ui*	‘onion’
	/zɔ̝n/	*zon*	‘sun’
	/hɔ̝n/	*kip*	‘chicken’
	/tɔ̝n/	*ton*	‘barrel’
Vfin (1 st person sing. present indic. *and* T imperative)			
	/ʀɛn/	*ren*	‘run’
	/wɪn/	*win*	‘win’
	/ɣ˧ʏn/	*gun*	‘grant’
ADJECTIVE			
	/dʏn/	*dun*	‘thin’
PROPER NAMES			
	An, Jan, Ben, Ien, Fien, Mien, Tien, Ton, Ron, John, all of which are first names.		

Except for a total of three verb forms (or rather, three pairs of homophonous verb forms), this set of non-deleting words consists of nouns and adjectives. As is obvious from (1), most n-deleting words are highly frequent in use.

Final n-deletion only applies to monosyllabic words, so polysyllabic words ending in a short vowel followed by the coronal nasal, such as /bəɣ˧ɪn/, *begin*, ‘begin’ (1 sg. pres. indic.; N), never undergo n-deletion. Moreover, variable final n-deletion only applies if the /n/ is preceded by a short vowel - so it does not affect monosyllabic words in which the final /n/ is preceded by a long vowel (such as /han/, *haan*, ‘rooster’) or a diphthong (e.g. /fɛin/, *fijn*, ‘nice’).[7]

In subsections 4.1 through 4.5 we will discuss this (quantitative) variation in n-deletion from a phonological point of view. In § 5 we will analyse the phonological and prosodic conditioning underlying the categorical blocking of n-deletion under certain conditions. In §§ 6 and 7 we will study the variability and change-in-progress in n-deletion with quantitative sociolinguistics methods. In the conclusion section 8 we will attempt to synthesize the phonological and sociolinguistic insights in the nature and dynamics of n-deletion.

4.1. Grammar or prosody?

The examples in (3) show that the deletion can also occur when the word at issue is the first member of a compound.[8] Put differently, n-deletion applies to prosodic words.

(3)

ɪɣ˧ɑŋk	<	[ɪn]$_W$ [ɣ˧ɑŋk]$_W$	*ingang*	‘entrance’, ‘entry’
vɑʔeː	<	[vɑn]$_W$ [eː]$_W$	*vaneen*	‘from one another’, ‘asunder’

One of the reviewers asked if the deletion cannot be accounted for along morphological lines (“delete /n/ before a morpheme boundary”). The components of a compound word are separated by a heavy morpheme boundary; lighter morpheme boundaries invariably involve flection. The only items in (1) which can undergo flexion are the verbs and the relevant flection endings are shwa, /s/ and /t/. Obviously, suffixing a shwa to a stem with final /n/ (which is only possible in the case of the verbs for ‘can’and ‘know’ 1st and 3rd person plur. pres.indic.) adds a syllable to the verb form, while suffixing /s/ and /t/ results in monosyllabic formations. -s only occurs in the case of the 2nd person sg pres.indic. verb forms for ‘can’and ‘know’, and -t only occurs in the case of the 1st person plur. pres. indic. verb forms for the verbs ‘go’, ‘get’, ‘have’, ‘do’, ‘say’, ‘see’, ‘stand’ and ‘beat’. In all these forms the /n/ cannot be deleted. In subsection 4.5 and section 5 below we will add further evidence to the claim that prosodic rather than grammatical constituency plays a role in n-deletion.

A rough formal description of the deletion process might look as follows:

(4)

$$/\mathrm{n}/ \quad \rightarrow \quad \emptyset \quad / \quad \overset{\sigma}{\overset{|\ \diagdown}{\begin{bmatrix}\text{-cons}\\ \text{+voc}\end{bmatrix}}} \;\_\_\;]_{W}$$

W = prosodic word

Clearly, some further specification is needed to prevent this rule from applying to words such as the ones in (2). Finding out the nature of this specification will be the main aim of the following subsections.

4.2. n-deletion and syllable structure

Generally, the right-hand margin of a syllable (or any higher-level prosodic constituent, for that matter) is a weak position for a consonant - witness the fact that assimilation usually occurs in that position. Yet, n-deletion violates an important constraint of languages such as Dutch, German and English, the constraint to the effect that syllables are not allowed to have a short/lax vowel in final position.[9]

Although German, Dutch and English have a strong dispreference for this type of output, there is *no complete ban* on it, as e.g. standard Dutch and several Dutch dialects (including the one spoken in Rimburg) have a few words of this very phonological make-up; all of these words (and there are some 3 or 4 of them in total) are interjections, however. Among the Rimburg ones are /bɑ/ 'ugh, yuck'and /wɑ/ (a tag), the standard Dutch ones include *hè* (a tag), *bah* 'ugh, yuck', *goh* 'geeh'.

Van Oostendorp (1995: 34) formalized this regularity in a projection constraint. Van Oostendorp's projection constraints relate the content of phonological features to prosodic structure - i.e. to words, feet, syllables, rhymes and nuclei. Van Oostendorp's constraint is reproduced here under (5).

(5) Connect(N', lax)= Project(N', lax) $\wedge$ Project(lax, N') $=^{def}=$ N$^{(0)}$ dominates [lax] *iff* N'

Put informally, this constraint says that lax vowels are only allowed in syllables with a branching rhyme, i.e. as part of a long/tense vowel or diphthong or a short vowel followed by a consonant. Independent of its formalization, this constraint is relevant to the output of the lexical component.

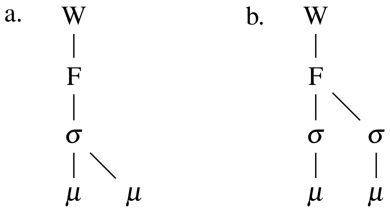
These considertations lead us to assume that in the dialects at issue in the items in (1) /n/ is present throughout the lexical component.

4.3. Lax vowels in stressed syllables in Dutch

If, however, structures with a short/lax vowel in open position are so heavily marked (as was pointed out in section 4.2 above), then why do all dialects of Dutch have words such as *lekker*, ‘good’, *binnen*, ‘inside’, *tussen*, ‘between’, *scharrel*, ‘rummage’, ‘deal’, ‘flirt’ (V pres.indic. 1 sg.), ‘flirt’ (N), etc., with a short/lax vowel in the stressed syllable?

Since Dutch stress assignment is quantity-sensitive, does this mean that the intervocalic consonants are ambisyllabic? If not, why do the syllables, despite the short vowel, apparently have enough weight to attract stress? Is the constraint against short/lax vowels in open syllables limited to word-final position? Or does the stress pattern of words of the type just mentioned fall out of Itô’s (1990) canonical minimal word, which consists of one foot (F)? The foot consists of one or two syllables, but invariably of two moras - vid.

(6)



The minimal word consists of one foot, which itself consists of two mora’s; these mora’s can either form one heavy syllable (6a) or be distributed over two light syllables (6b). In the second case, the left-hand syllable will head the foot. Bisyllabic words with a short//lax vowel in the first syllable (such as *lekker, binnen, tussen, scharrel*) fit the template in (6b).

In the next subsection we will consider the question if in the dialects at issue the items in (1) fit the template in (6a), i.e. if the /n/ is indeed present, or not.

4.4. n-deletion or n-insertion?

There are several facts and considerations regarding the words in (1) that point towards a deletion process ($C_0Vn \rightarrow C_0V$) rather than an insertion process ($C_0V \rightarrow C_0Vn$):

(i) the set of words involved is limited. Native speakers never overgeneralize the application of the rule to any of the few words with a short/lax vowel in final position, i.e. no vowel-final words ever get an /n/ epenthesized hypercorrectly;
(ii) dialects which do not have this deletion rule, invariably have final /n/ in the words in this set - although dialects can differ in their lexical representations (a possibility which was considered "rare" by Chomsky and Halle 1968: ix-x);
(iii) if the /n/ were not present at least lexically, then neither from the point of view of weight sensitive stress assignment nor from the point of view of prosodic minimality (cf. (6a)) would it be explainable why the words in this set can be stressed *even* in case the /n/ is not phonetically present (cf. Weijnen 1991: 149). The following 'real life' examples, taken from the corpus of conversational dialect use that was recorded for this study (see section 6 below), serve to illustrate this.

(7)

a. \`ɪdʏijə
induwen 'push in'
b. \`lɛgət \`doːᵊ mɑR \`ɪ
leg 't daar maar in 'put it in there'
c. wɑt \`wølstə ɣ˧eːᵊR \`hɑ
wat zou je graag willen hebben? 'what would you like to have?'

In these examples the relevant n-less forms bear word stress -as in (7a)- or even intonational stress - as in (b) and (c). Since Dutch stress assignment is weight sensitive, /n/ must be lexically present in these words.

What exactly is it that distinguishes the words that can undergo final n-deletion, i.e. the words in (1), from the words that cannot? In the next subsection we will establish that there are several for properties which set the items in (1) apart from words of the type in (2).

4.5. Enabling causes

Apart from the fact that they are frequent to highly frequent in use, what all the n-deleting words have in common is:
(i) the structure of their rhyme;
(ii) the fact that they may but need not attract a certain degree of stress - as illustrated in (7a-c) above;
(iii) the fact that they may but need not be realized with a tone contour.
Ripuarian and Limburg dialects of Dutch are pitch-accent languages.[10] They have two different tone contours. One is HH (phonetically HLH; 'circonflex') and the other is HL ('acute'); these two contours are both morphologically and lexically contrastive, witness minimal pairs such as:

(8)

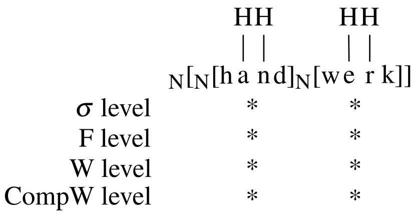
a.	ɛ̌ʀᵊm	:	ɛ̀ʀᵊm	'arm' : 'arms'
	dǎx	:	dàx	'day' : 'days'
b.	dǔ	:	dù	'you' (2 sg. nom.) : 'then' (adv.)
	mǔᵊʀ	:	mùᵊʀ	'carrot' : 'wall'

In the analysis proposed by Hermans (1984; 1994), words with the HH contour have a High tone in their lexical representation, while words with a HL contour do not have any tone in their lexical representation.

For a syllable to be realized with a tone contour at all, two conditions need to be met: a segmental and a supra-segmental one. The segmental condition requires there to be a vowel and a subsequent sonorant in the rhyme. Hence the rhyme can consist of any of the following: a short vowel followed by a sonorant consonant, a long vowel or a diphthong. The suprasegmental condition is that the syllable at issue has primary or secondary stress, so e.g. compound words[11] the two component parts of which contain syllables which meet the segmental conditions have two contour tones. Vid. for *handwerk*, 'handiwork', 'fancywork', 'manual work':

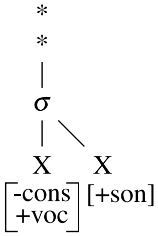
(8)

c.



Formally, syllables that can be realized with a tone contour can hence be represented as in (9).

(9)



If sufficiently stressed, the words in (10), which constitute a proper subset of those in (1), can be realized with either tone contour HH or HL.

(10)

PREPOSITIONS			
	/vɑn/	*van*	‘of’, ‘from’
	/ɪn/	*in*	‘in’
ADVERBS			
	/dɑn/	*dan*	‘then’, ‘than’
	/nun/	*nu, nou*	‘now’
Vinfin			
	/hɑːn/	*hebben*	‘to have’

All the other words in (1) can only have the tone contour HL - provided the metrical condition is met.

Even if sufficiently stressed, no tone contour occurs on any of the words in (1) if the /n/ is deleted. It follows that in these words /n/ is a tone bearing unit;

(iv) with the exception of the conjunction *en*, ‘and’, all words in (1) may occur in sentence-final position, which (being Intonational Phrase or even Utterance-final) is usually a prosodically weak position;
(v) the prepositions, the conjunction, the auxiliaries and the modal verbs are all function words, which explains their highly variable stress, pointed out in (ii) above (cf. Inkelas and Zec 1995: 544);
(vi) with the exception of the verbs, none of the words in (1) is a lexical head - so they never head a phonological phrase. In ordinary spoken Dutch, auxiliaries and modal verbs are often immediately followed by an infinitive or a past participle; in such a constellation, phonological phrase restructuring (Nespor and Vogel 1986: 173) usually makes the auxiliary / modal lose its status as a phonological phrase of its own. Hence of the relevant prepositions, conjunction, adverbs, auxiliaries and modal verbs, some are always and the others are usually in a prosodically weak position in that they do not head a phonological phrase.

It appears that the combination of

a. their usually being in a prosodically weak position,
b. the highly variable behavior of these words with respect to stress, and
c. the fact that if unstressed, these words cannot be realized with a tone contour, in which case /n/ loses its function as a tone bearing unit,

plays a role in the explanation of the fact that in these words final /n/ can undergo deletion - and, *mutatis mutandis*, of the fact that words such as the ones in (2) never undergo final n-deletion.

A problem for this account is the fact that the athematic verbs in (1), as well as /kɪ̞n/, ‘(I) know’and the infinitival form /hɑːn/, ‘to have’, all of which equally head a phonological phrase, are nevertheless input to variable n-deletion. One can think of a scenario in which the athematic verbs, /kɪ̞n/, and the infinitival form form /hɑːn/ became part of the input to the rule at a later point in time. The extension of n-deletion to these words would then be a matter of analogy, *in casu* lexical diffusion, or lexicalisation. In the latter case, the thematic verbs, /kɪ̞n/, and /hɑːn/ may be ‘precompiled’[12] for variable n-deletion (Hayes 1990), on the understanding that (1) the syntactic environment, which in Hayes’ proposal is referred to by means of ‘frames’, does not play a role, and (2) also in these words n-deletion applies *variably*, despite the fact that precompiled rules are lexical.

The present account of the motivation of n-deletion is thus a genetic explanation; this account is partly comparable in nature to that of the well-nigh categorical post-shwa word-final n-deletion in most varieties of modern Dutch (including the standard variety). Since this n-deletion, which is not confined to monosyllabic words, is not sensitive to the morphological status of /ən#/ (Van Oss and Gussenhoven 1984), it is probably a postlexical rule.

Kiparsky (MS) drew a distinction between effective and enabling causes of processes of language change. While effective causes are "structural preferences grounded in UG", enabling causes are specific internal and/or sociolinguistic factors operating in a given situation. The present account of dialectal n-deletion is based on a cluster of enabling causes or, more precisely, conditions favoring n-deletion, which reinforce each other's effects.

As regards the multitude and complexity of its internal conditioning, n-deletion resembles French liaison (as described by, e.g., Booij and De Jong 1987).

5. Two conditions blocking n-deletion

Dialectal n-deletion, though variable, is subject to two blocking conditions. The first of these can be seen at work in (11a-e). While n-deletion is permitted in (11a), it is not in (11 b-e), where the /n/ equally seems to be in word-final position.

(11)

a.	dɑt	kɑ(n)	ɪç
	'that	can	I'

b.	ɪç	kɑ*(n)	ət
	'I	can	it(clit)'

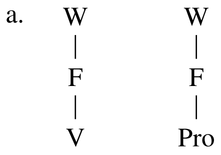
c.	ɪç	hɑ*(n)	əs	ɔ̝x	ɑl	ɣ˧əhɑt	di ...
	'I	have	ones(clit)	also	already	had	that', i.e. 'I have also already had ones that ...'

d.	ɪç	hɑ*(n)	ə huːs	ɪ	kɪRəkRoə
	'I	have	a(clit) house	in	Kerkrade'

e.	kɑ*(n)	ə	koːmə
	'can	he(clit)	come?'

The difference between shwa and other vowels in the right-hand context suggests that the n-deletion rule is blocked because the following word cliticizes (hence the reduced vowel) and as words in Dutch cannot have a schwa initially it gets incorporated into the target word. Whereas the prosodic structure of the relevant part of (11a) is[13]

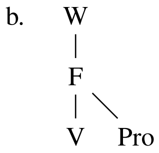
(12)



n-deletion okay

in the case of (11b) through (11e), where n-deletion is blocked, it is

(12)



n-deletion blocked

The /n/, it appears, can only be deleted in word-final position. The relevant blocking constraint falls out automatically from the formulation of the rule in (4) above.

Independent of the question whether in environments in which n-deletion is blocked as a result of enclificìzation, the /n/ becomes ambisyllabic, as in (13a), or gets resyllabified, as in (13b), it will not be in word-final position, so it cannot fall prey to the deletion process.

(13)

Aside: there is an interesting parallel between the blocking of the dialectal n-deletion in this non-final context and a variable process of 'euphonic' n-insertion which operates in many varieties of *spoken Dutch*, including colloquial standard speech. The hiatus between a shwa-final word, suffix or clitic and a subsequent shwa-initial enclitic forms the context of insertion of the 'euphonic' /n/. Examples are:

(14)

clitic - clitic:
a. *Daar heb je ‘m* ~ *Daar heb je-n-’m*
[jəʔəm] [jənəm]
There have you(clit) him(clit)
‘There you have him’

suffix - clitic:
b. *Ik hoorde ‘m* ~ *Ik hoorde-n-‘m*
[hoəːᵊʀdəʔəm] [hoəːᵊʀdənəm]
I heard him(clit)
‘I heard him’

shwa-final word - clitic:
c. *Als we de schade ‘s schatten* ~ *schade-n-‘s*
[sxadəʔəs] [sxadənəs]
If we the loss(shwa-final) just(clit) estimate
‘If we just estimate the loss’

d. *Toen gaf Abe ‘m een lel* ~ *Abe-n-‘m*
[abəʔəm] [abənəm]
Then gave [abə] him(clit) a hefty kick
‘Then Abe gave him a hefty kick’

(15)

but
e. *Toen gaf Laura ‘m een zoen* ~ **Laura-n-‘m*
[lɑuraʔəm] [lɑuranəm]
Then gave Laura him(clit) a kiss
‘Then Laura gave him a kiss’

f. *Daar heb je Arie* ~ **heb je-n-Arie*
[hɛpjəʔari] [hɛpjənari]
There have you Arie
‘There’s Arie’

As (15e) and (f) show, the insertion only takes place if the word ends in the reduced vowel and if a shwa-initial clitic follows. The prosodic words resulting from the double encliticization exemplified in (14a-d) constitute ternary feet.

Whereas n-deletion is *categorically* blocked whenever a shwa-initial clitic follows, the rule is *sometimes* prevented from operating in case the following word has an initial /d/; the rest of the segmental or suprasegmental structure of the following word does not play a role in this connection.

The fact that the rule is never blocked if the following word has, say, /t/, /s/ or /z/ as its initial segment makes it easy to understand the segmental part of this second constraint: the shared features [+voice], [-cont] and [cor], and hence the partial gemination of /n/ and /d/ apparently have something to do with it. *Feature sharing* on both the laryngeal and the (supralaryngeal) manner and place tiers thus forms the essence of the segmental part of this blocking constraint.[14] In a structure of this type, n-deletion would lead to a violation of the inalterability (or integrity) of geminates (Hayes 1986; Kenstowicz 1994: 410-17).

However, this does not explain why n-deletion is not always blocked before a d-initial word. The apparent unpredictability of blocking if a d-initial word follows, in combination with the fact that n-deletion as such is already a variable rule, makes it difficult to further elucidate the linguistic motivation for this blocking. It would not be unreasonable to suspect that a prosodic factor is at work here as well. The first step to test this hypothesis is to delimit the domain of this specific constraint by determining in which cases n-deletion may take place if a d-initial word follows.

As was already clear, in general n-deletion is never blocked if the relevant word is the first member of a compound (cf. (3) and (7a) above). It was exactly for this reason that it has been concluded (in section 4 above) that final n-deletion operates at the level of the prosodic word, and is therefore a domain limit rule (Selkirk 1980; Nespor and Vogel 1986). It may operate at the juncture of two intonational phrases, as is evidenced by the following example - which, like all other examples given in this section, is taken from the spontaneous dialect use recorded for this study.

(16)

mɑR	ɪç	sɑ(n)	dɑt	ɪs . . .
‘but	I	say:	″that	is . . .’

These facts suggest that the apparently coincidental blocking of the n-deletion rule in case the initial segment in the following word is /d/, has the prosodic level in between the phonological word and the intonational phrase as its domain; this is the level of the phonological phrase (Φ).

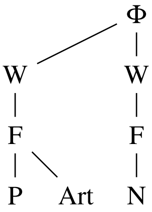
Having determined the prosodic domain, the next step will be to consider more examples in order to establish the exact nature of the constraint. Final n-deletion is not allowed and never occurs in:

(17)

a.	vɑ*(n)	də(ʀ) . . .
	‘of	the . . . ’
b.	ɪ*(n)	də(ʀ) . . .
	‘in	the . . .

Here again enclitics are involved - in this case the cliticized forms of articles. The prosodic structure of sequences such as the ones in (16) then is:[15]

(18)



n-deletion blocked

Here /n/ is not in word-final position, and can therefore not fall prey to the deletion process. Articles need not be realized in a phonetically drastically reduced fashion; but even when they are not, n-deletion is still blocked.

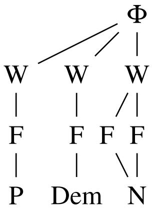
In the following example, the d-initial word is a demonstrative. Phonetically it is not reduced (in this specific realization it was even heavily stressed, which is why it has been underlined), but n-deletion is blocked all the same:

(19)

ɪ*(n)	dɑt	filiʲal
‘in	*that*	branch’

In this case, the prosodic structure is:

(20)



n-deletion blocked

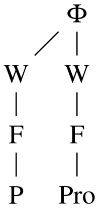
Had the demonstrative not been stressed, n-deletion would also have been blocked. Cases where n-deletion is allowed before a d-initial word include:

(21)

a.	ɑ(n)	dɛm
	'on/to	him'
b.	vɑ(n)	dɪç
	'of	you'

Assuming that personal pronouns count as lexical heads,[16] the prosodic structure of such configurations is:

(22)



n-deletion okay

However, in other cases where the d-initial word is a pronoun, the final n-deletion rule is blocked, as for example in:

(23)

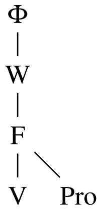
ɪç	kɑ*(n)	dɪç	nɪt	hɛlᵊpə
I	can	you	not	help

'I cannot help you'

Here [kɑn] bears constrastive stress; the following pronoun is unstressed, and encliticizes (cf. Neijt 1985: 186):

(24)

a.

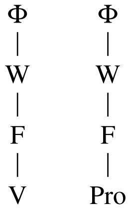


n-deletion blocked

Here, again, /n/ is not in word-final position and can therefore not be deleted. Had the following pronoun been stressed, n-deletion would not have been blocked. In case the pronoun is stressed, it is a prosodic word of its own and projects a phonological phrase:

(23)

b.



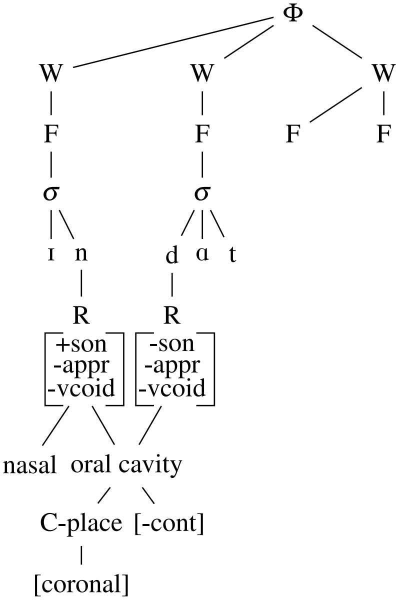
n-deletion okay

The solution to the problem regarding the non-segmental part of the second blocking constraint on the deletion of /n/ in word-final position thus is:

(24) In case the following word has an initial /d/, final n-deletion is allowed if and only if the following word is a lexical head, thus heading a phonological phrase.

The nature of this constraint appears to be consistent with the fact -uncovered earlier in this section- that the domain of the deletion process is the phonological phrase. Whenever the neighbouring d-initial word does not project a phonological phrase of its own, /n/ and /d/ obviously geminate partially. The relevant part of an example such as (18) is hence structured as follows:[17]

(25)



Clearly, in this type of constellation n-deletion would violate the principle of the inalterability of geminates - and of multiply linked structures (Kenstowicz 1994: 218) generally. It can hardly be a coincidence that word-final n-deletion in the historically related *Moselfränkische dialects* (*in casu* Lëtzebuergesch) studied by Gilles, while different at the surface, turns out to be subject to exactly the same blocking (2006: 52).

For a constraint-based account of the insights gained so far into both the complex, multiply conditioned variability in n-deletion and the blocking of the process, see Hinskens (in preparation).

There are two reasons why is it important to know that there are blocking conditions, and what the nature of those blockings is. First, it leads to a deeper insight into the nature of the deletion process itself. The second reason is more practical in nature: one should not look for variation in the application of a rule where there is none, i.e. where the rule is not allowed to apply. Doing so would have an invalidating effect on the results of the statistical analyses of the variance in n-deletion, basically compressing the figures. After all, the two conditions are ‘knock-out factors’, in the variationist terminology (Labov 1969; Rousseau and Sankoff 1978).

Section 7 zooms in on both linguistic and extra-linguistic determinants of the main patterns in the quantitative variation in n-deletion in the Rimburg dialect. Section 6 introduces the method underlying this part of the study.

6. Tracing dialect levelling. Methodological aspects and considerations

Word-final n-deletion is one of the dialect features which were studied in the Rimburg dialect levelling project. For the project at large, dialect use was recorded in a sample of 27 native speakers of the dialect. Most of these people had been born and bred in Rimburg, all of them had lived in Rimburg from -at the most- their 6th year of life onwards, and at least one of their parents were native Rimburgers. The sample of speakers was stratified for age group. Three different age groups were represented in the sample, the age intervals being 20-30, 40-50 and 60-75. These three groups are referred to as Younger, Middle and Older, respectively; each age group was represented by 9 speakers. The methodology for this project is based on the assumption that dialect levelling may manifest itself as a statistically significant decrease in the use of dialect features from the Older through the Middle to the Younger generations of speakers - the so-called apparent time method. The sample of speakers was homogenized for sex, in that only men were included. Within margins, the sample was also homogenized for socio-economic (SE) background. SE background was operationalized through two variables: educational background and occupational level. Both variables were measured on six-point scales. The homogenization was reached by including only people whose score on the mean of these two scales is minimally 2 and maximally 4. So in this respect there is variation within the sample - the variation spanning exactly half of the scale's range. All averages of the SE background distribution are to be found in this part of the scale: median, mode and mean.

For the sake of collecting the speech material, of every speaker in the sample recordings were made of three types of dialect use. To obtain a sufficient number of realizations of each single dialect feature by each single speaker, dialect use was elicited through a wide variety of tasks. Relatively spontaneous, conversational dialect use was recorded in two different types of situation: in-group and out-group contact. Conversations were organized in three groups of each three members of each age group; these are referred to as in-group conversations. The out-group conversations consisted of each individual speaker talking to a speaker of a different dialect.

Distilling the relevant data from the recorded speech material involved analyses of various types. The analyses consisted of several steps. The end result consisted of figures representing each speaker's relative use of each dialect feature - both on the overall level and in a range of linguistic conditions. Ultimately the analyses were quantitative in nature; the end result were indexes ranging between 0 and 100. (Cf. Hinskens 1992: Ch. 4 for an in-depth account of the considerations underlying the decisions taken at the various steps in the methodology.)

Section 7will zoom in on the findings from quantitative analyses of the n-deletion data distilled from the recorded in-group conversations. The occurrence or non-occurrence of n-deletion was not studied in case a word follows with a nasal in initial position nor, of course, in the constellations in which n-deletion is blocked (discussed in the preceding section). The statistical analyses are based on approximately 900 observations for the entire sample of speakers, which comes down to an average of 33 tokens per speaker.

7. n-deletion in relation to several linguistic and extralinguistic parameters

In this section we will first analyse the quantitative variation in n-deletion in the entire sample of speakers (7.1.) and then we will zoom in on the apparent time dynamics of the process in the speech community (7.2).

7.1. Variation in n-deletion

For the sake of this study the relative occurrence of n-deletion was related to five different linguistic dimensions (or 'factor groups'). The findings for the effects of each linguistic dimension on n-deletion are presented in Table 1; in this table, '–' signifies the absence, and '+' the presence of a statistically significant effect.

The gradience in the deletion process appears to be directional in several dimensions. The (highly significant) effect of the right-hand environment is brought about by the fact that n-deletion applies most frequently before a consonant-initial word, least before a vowel-initial word, with a following pause, i.e. any major prosodic break, triggering intermediate deletion rates. The relatively high occurrence of n-deletion before consonants is not surprising in view of the fact that the output, C_0V \$ C_1V, satisfies the universally

Table 1: The use of the n-deletion rule in several linguistic conditions; means and the significance of the differences between the means.

+	before consonant 60.10 > pause 56.06 > vowel 21.29 F=21.37 df=2,42 *p=.000*
−	ɪ-lowering relevant 41.52 - not relevant 46.44 t=-.95 df=26 p=.349
+	in adverbs 59.97 > prep. 48.14 > verb forms 28.30 F=24.92 df=2,52 *p=.000*
+	stressed 33.81< unstressed 59.99 t=-5.09 df=26 *p=.000*
+	lexical H 61.43 > no lexical H 25.64 t=8.96 df=26 *p=.000*

preferred syllable sequencing as well as -more specifically- the syllable contact law (Murray and Vennemann 1983; Vennemann 1988). The relatively low proportion of n-deletion before vowels can be equally explained by general phonological preferences and dispreferences - in this case of the general disliking of hiatuses. The relatively high deletion rate before pauses is remarkable, however, as the output is heavily marked: a short vowel not only in syllable-final position, but at the right-hand edge of a higher prosodic constituent - commonly at least a phonological phrase.

The sensitivity of the n-deletion process to the phonological nature of the right-hand environment suggest that it is a *postlexical* rule. The ban on syllables with a short/lax vowel in open position (formally represented in (5) above) is a lexical constraint. If n-deletion is indeed a postlexical process, then derivationally it does not violate this constraint.

As far as the segmental environment goes, attention was also paid to the potential interaction between n-deletion and the variable dialectal lowering of the short unrounded high front vowel /ɪ/ before nonlabial nasals. Depending on the rule-typological status of the two processes, n-deletion may bleed ɪ-lowering. In order to bring to light any effects of the relevance or irrelevance of ɪ-lowering on n-deletion, the set of words relevant to n-deletion was divided into two subsets: the ones with and the ones without the short unrounded front vowel. The difference in proportion of n-deletion between the two sets and hence the influence of this factor is not significant, however.

For the analysis of the effect of the factor group ‘grammatical class’, i.e. of the syntactic category membership of the potentially n-deleting words, the conjunction *en*, /ɛn/, ‘and’, was left out of the analyses, since the examination of the corpus of elicited dialect use by the same speakers had made clear that this conjunction shows relatively little n-deletion, as appears from the overall sample average of 7 / 100. The in-group conversational dialect use by the Older speakers contained no more than 4 / 78 cases of n-deletion in *en*. Logically, these 4 instances of deletion were produced by no more than four different speakers. Actually, only 3 out of the 9 Older speakers occasionaly delete the /n/ in this conjunction; even for these few speakers, there is very little variability in the deletion of the final /n/ in this very frequently occurring conjunction. In all, it appears that the conjunction is not a core member of the potential for the deletion rule. The factor group grammatical class was therefore confined to adverbs, prepositions and verbs.

Table 1 shows that the grammatical class has a significant effect; adverbs show most, verb forms show least deletion; prepositions have intermediate rates of deletion. However, it is very uncommon for postlexical rules to be sensitive to grammatical information; it is much more likely for them to be affected by prosodic factors and stress. Of the grammatical classes represented by the n-deleting words, verbs are the only ones which constitute lexical heads and can thus head phonological phrases. At average they will probably attract relatively large amounts of stress. As can be seen in Table 1, stress (which was implemented as the ability to form an Abercrombian foot, i.e. the phonetic type of foot that may consist of syllables of adjacent words) does as such have a significant effect on n-deletion - cf. (36) above. Table 2 shows that what looks like a grammatical effect may well be a prosodic effect.

Table 2: Pearson correlation coefficients between several linguistic conditions as regards n-deletion (two-tailed probabilities).

	unstressed	stressed
adverbs	.3438 (.079)	*.5657* (.002)
prepositions	*.8141* (.000)	-.0010 (.996)
verb forms	.1624 (.418)	*.8288* (.000)

As far as the relative frequency of n-deletion is concerned, there appears to be a strong linear relationship between prepositions and not being stressed

(r^2=.66) on the one hand, and between verb forms and being stressed (r^2=.69) on the other. This finding seems to support the conjecture that prosodic factors and stress form the genetic explanation for n-deletion (cf. section 4.5 above).

The factor group 'tone contour', i.e. the presence or absence of a floating High tone in the lexical representation, also turns out to have a significant effect on n-deletion, in that the forms that have a floating High tone, i.e. the ones that can be realized with a HH tone contour, i.e. the words belonging to the subset in (10), undergo significantly more n-deletion than the words which do not have a High tone. But again, as n-deletion is probably a postlexical rule, it is not likely to be sensitive to lexical information such as the underlying presence or absence of a floating High tone. Closer study on the basis of independent data (De Jong 1979, based on the corpus of spoken contemporary Dutch, collected by Heikens in Amsterdam) suggests that the effect of the factor tone contour on n-deletion may well be a token frequency effect. In general, the standard language variants of the HH words in (1), i.e. the subset of the words which have a floating High tone, occur about ten times as frequently as the words which in the dialect can surface only with a HL contour, i.e. the words which do not have a High tone in their lexical representation (see Hinskens 1992: 363-64 for details). Hence, the different degrees of n-deletion in both sets could be indicative of different degrees of wear and tear, so to speak.[18]

7.2. Levelling out of n-deletion

Breaking down the sample average for n-deletion for the factor age group renders the pattern in Table 3. The apparent time decrease in n-deletion turns out to be almost significant at the 5 % level.

Table 3: Analysis of variance: mean and standard deviation of the use of the n-deletion rule in the three age groups; the probability and the explained variance.

age group	*	mean	s	F	df	p	%var
O	9	53.08	11.55	3.184	2,24	.0594	9.84
M	9	38.95	5.16				
Y	9	43.14	16.92				
entire sample	27	45.06	13.18				

* n of speakers

Whereas there is no clear age effect on the overall level, n-deletion shows significant age group effects in four specific linguistic conditions. These are presented in Table 4, along with the significance of the interaction effect of age group and the linguistic factors involved.

Table 4: Analyses of variance. Significant effects of the variable age group on the use of the n-deletion rule in several linguistic conditions. The significance of the effect of the interaction between age group and linguistic dimension.

	F	df	p	%var	*
I-lowering not relevant	3.438	2,24	.0486	9.30	–
adverbs	7.934	2,24	.0023	25.31	–
with lexical H	4.914	2,24	.0163	13.64	–
before consonants	4.459	2,24	.0226	22.98	+

* signif. interaction effect age gr. x ling.dim.?

As pointed out in the discussion of the findings in Table 1, three of these four factor groups significantly affect the overall sample rate of n-deletion as such; ɪ-lowering does not.

In words which do not contain an /ɪ/, in adverbs and in words with a floating High tone in their lexical representation or a high average frequency of usage, the age group effect consists of a significant decrease in the application of the n-deletion rule between the speakers of the Older age group on the one hand and the representatives of the Middle and Younger age groups on the other. In these conditions there is no significant difference in n-deletion rates between the Middle generation and the Younger speakers. The ideal-typical gradual pattern of loss Older > Middle > Younger only occurs in the application of the n-deletion rule before consonants. This last condition is part of the only linguistic dimension (*in casu* 'right-hand environment') which turns out to significantly affect the levelling process, as appears from the effect of the interaction between this linguistic dimension and the speakers' age group (F=2.65 df=4,38 p=.048). The n-deletion rates increase before pauses but diminish before consonants and vowels. Obviously, the deletion rule is under pressure where it creates the least (_C) and less (_V) marked output, while it flourishes where it creates the most (_pause) marked output. So the directionality of the levelling process in this dimension does *not* seem to have an *internal* motivation.

After trading in the three-value variable age group for the ratio variable age in years, the part of the variance in n-deletion in the overall n-deletion climbs from 9.8 (Table 3) to 11.9. This figure results from a linear regression analysis with the speakers' age in years as a predictor for n-deletion on the overall level (b=.244, β=.345, t=1.839, p=.078). Obviously, the speakers' age in years accounts for an additional 2.1 percentage points of the variation in n-deletion.

These findings very clearly point to the fact that there is no direct deterministic relationship between the speakers' ages and the rate of n-deletion. Generally, the speakers' age or age group cannot as such explain any processes of linguistic change such as dialect levelling. Significant age (group) effects on language use need to be related to societal developments which in turn have sociolinguistic effects. In connection with these societal developments underlying processes of linguistic change cf. Mattheier's (1980: 140-58) 'allgemeine gesellschaftliche Modernisierungsprozesse', i.e. general processes of societal modernization, and Van Hout's (1989: 273-77, 323-24) 'societal developments'. Such processes are probably more directly reflected in speaker background variables other than age or age group.

In the Rimburg project two bundles of speaker background variables were studied in order to unearth some of the societal changes underlying changes in dialect use: geographical background and social background. The geographical background variables are autochthony (the degree of 'nativeness' of the speakers in the sample to Rimburg) and geographical mobility; the factor geographical mobility is inversely related to what Mattheier (1980) has labelled 'Ortsloyalität', loyalty to the village. The speakers' social backgrounds are assessed through their educational background on the one hand and their occupational level on the other. All four background variables were measured on ordinal scales (see Hinskens 1992: § 4.3.3 for details).

As far as dialect use is concerned, the interrelatedness of the speakers' ages and the four macro-social speaker background variables can be traced in two ways. The first approach is based on the assumption that there is a triangular connection between age, speaker background and dialect use. This has been visualized in Fig. 1.

Figure 1: Modelling the interrelatedness of the speakers' age, background and dialect use.

All three relations were studied quantitatively:
– speakers' age - speakers' macro-social background
– macro-social background - dialect use
– age - dialect use

Significant relations were found in each of the three respects (see Hinskens 1992: §§ 4.3.3, 9.2.1). In order to separate the age effect from the effects exercized by the four macro-social variables, correlation coefficients were calculated for the relationships between age and n-deletion both before (r=.3451, p=.039) and after partialling out the effect of the four background variables (r=.3011, p=.087). Comparison of the coefficients and their significance shows that a part of the age effect on n-deletion is clearly due to variation in the speaker background variables - though not a large part: $(.3451)^2$ - $(.3011)^2$ = .0284. In other words, no more than approximately 3 % of the variation in the overall n-deletion rate is related to the four macro-social speaker background variables autochthony, geographical mobility, educational background and occupational level.

The second way to trace the relationships between age and spreaker background variables is to weigh them. If societal developments are reflected in macro-social speaker background variables, then the speakers' age should be outweighed by their macro-social background variables. To this end a stepwise multiple regression analysis was carried out with years of age and the four speaker background variables as predictors of n-deletion. The one and only predictor which remains in the final equation is geographical mobility: b=-1.207, β=-.421 t=-2.277 p=.032 MR=.421 %var=17.8 (F=5.18 df=1,24 p=.032).

The outcomes of this analysis show that the age effect on n-deletion solely consists of the growing geographical mobility of the dialect speakers. The more mobile the speakers get, the more they will be in contact with speakers of different dialects, the less they tend to delete postvocalic final /n/'s in monosyllabic words.

So much for n-deletion on the general, overall level. For the linguistic conditions in which the n-deletion rule is significantly losing ground (Table 4 above), identical analyses were carried out to determine the interrelatedness of the speakers' age and the four macro-social speaker background variables. These analyses were expected to provide an insight into the possible interplay between linguistic and social forces in what may be the initial stages of a process of rule loss. Consider the results in Table 5.

Table 5: Findings from stepwise multiple regression analysis for years of age and the four speaker background variables. Criteria: n-deletion in linguistic conditions showing loss. Predictors within the .05 probability level for variables to be entered.

criterium	predictor	b	β	t	p	%var
ɪ-low. not relevant	—					
adverbs	years of age	.442	.490	2.754	.011	24.0
with lex H	education	-.062	-.426	-2.308	.030	18.2
before cons	education	-.079	-.480	-2.682	.013	23.1

Not even years of age is entered into the equation for n-deletion in words in which the ɪ-lowering rule is not relevant. This means that the age effect must be rather weak there, so that after all in this condition the rule is perhaps not really endangered (or is not endangered as yet, one might speculate).

Age appears to be the only predictor which has a clear effect on n-deletion merely in adverbs. As was pointed out above, there is a sharp decrease in n-deletion in adverbs after the Older age group, such that the speakers of the Middle and Younger groups form one homogeneous subset, even though the Younger speakers use the rule slightly more often.

The decreasing n-deletion in words with a lexical High tone (and at the same time a relatively high token frequency) and before consonants can be predicted from the speakers' educational background - in the negative sense: the decreasing application of the n-deletion rule in words with a lexical High tone and in case a consonant-initial word follows is a consequence (for over 18% and 23% respectively)[19] of the general increase in education in the first place. In the Dutch society, this general increase in education is a post-World War II development. Note that with education comes literacy and familiarity with the orthography, including the written forms of the words in (1). Literacy and writing are limited to the standard language, as the dialect is not in written use. Moreover, the dialect hardly plays a role at school as a medium of instruction (cf. Vallen and Stijnen 1996), let alone as a subject.

The finding for the factors underlying the decreasing n-deletion in words with a lexical High tone (or high average frequency of usage) and before consonants could mean that the rule is gradually changing from a traditional dialect feature into a linguistic variable with social meaningfulness. In the terminology of Coseriu's (1987; 1988a) 'Architektur der historischen Sprache': in certain linguistic conditions this dialect feature may be undergoing a change from diatopical into diastratal variation. This may explain why n-deletion seems to hesitate between loss and maintenance. What remains of the rule may (at least partly) be a sociolect feature, rather than a dialect feature.

8. Conclusion and discussion

The findings regarding the apparent time dynamics in the n-deletion process point toward a change in the nature of this dialect feature. The outcomes of the multiple regression analyses suggest that n-deletion may be developing from a merely geographically into a socially or even socio-stylistically meaningful feature, i.e. from an *indicator* into a *marker* (Labov 1972).[20]

In most urban communities in the Old World, the features of the traditional dialects which survived processes of dialect levelling have generally changed from indicators into markers (for references cf. Auer and Hinskens 1996, § 3.1). It is fairly common for features defining urban dialects to be clearly marked both socially and stylistically. Some of these features are now so heavily stigmatized that they have reached the *stereotype* status. For a rural dialect like the one spoken in Rimburg, however, the type of change which leads to a dialect feature attaining socio-stylistic meaning seems to be a new type of development. This may be related to the demographic restructuring which resulted from the industrialization of the region in the first decades of this century (briefly sketched in section 3 above).

This type of change in the sociolinguistic status of a dialect feature sheds a new light on an insight which Edward Sapir expressed as follows: “Old dialects are being continually wiped out only to make room for new ones” (1921: 152). Applying this claim to the above findings regarding word-final n-deletion, it can be established that

(1) in certain linguistic conditions the rule is being levelled out. This affects variation between the Rimburg dialect on the one hand and the non-n-deleting Limburg and other dialects of Dutch, *including the standard variety*, on the other. So the (partial) loss of n-deletion is not an instance of koineization (cf. section 2 above; see also Siegel 1985; Hinskens 2001);

(2) in some linguistic conditions the application of the rule is no longer geographically but socioeconomically meaningful. So it appears that, rather than being completely ousted from the grammar and the larger verbal repertoire, in a few linguistic conditions n-deletion will in the future be recycled.

More generally, Sapir's claim is supported by the fact that, throughout the Old World, traditional local dialects are being gradually replaced by new, relatively homogeneous regional dialects and/or by regional varieties of the standard language. Findings for other Rimburg dialect features (Hinskens 1992) strongly point in this direction as well.

The study of language variation and change on the one hand and formal linguistic theory on the other can both profit from a further integration (Coseriu's 'internal interdisciplinarity' - 1988b). This has been illustrated by the analysis of the dialectal n-deletion rule, specifically for two aspects:

– the very existence of the phenomenon raises a number of questions regarding the exact nature of the alleged lexical output constraint active in a group of Germanic languages which disfavours short/lax vowels in open position. Is there maybe more to it? (§ 4.2 and 4.3);
– on the other hand, phonological theory, in particular a theory of the hierarchical organization of distinctive features and a theory on the interface between phonology and syntax, has served to disentangle and formally explain the nature of the conditions under which n-deletion is blocked. The outcomes of these analyses, in turn, further clarified both segmental and prosodic aspects of the motivation for n-deletion (§ 5).

In general it appears that the chances for a cross-fertilization between sociolinguistics and formal linguistic theory to be successful will be better if those who work in formal theory are willing to free themselves of some of the Saussurian chains in which they have bound their approaches, in particular from such fundamental distinctions as those between *langue* and *parole* and *synchrony vs. diachrony*. These two analytical distinctions probably strengthened generations of linguists in their conviction that processes of language change cannot as such be observed and studied. In reality, *langue* changes gradually in *parole* - and it is in *parole* that older and newer states of a *langue* coexist synchronically as language variation. The Saussurian distinctions have proven to be extremely useful for analytical purposes, but meanwhile a situation seems to have been reached in which rigidly clinging to these analytical tools impedes progress in the understanding of language change (cf. Guy 1997). It has been pointed out that formal theory cannot be required to account for diachronic facts or processes of language change. However, in a situation in which two theories perform equally well in dealing with the relevant synchronic data, the one which can also account for diachronic facts, is the more adequate theory (cf. Kiparsky 1971). In formal theory, more weight should be given to this evaluation metric.

At the same time, however, neither area of linguistics should abandon the Saussurian principle that, generally, the relationship between linguistic form and grammatical function or semantic 'content' is entirely *arbitrary*. This insight extends to language variation and linguistic diversity generally - cf. "We find in principle and in fact that certain differences do not make a difference" (Labov 1972). The opposite standpoint (represented by Coseriu; see also e.g. García 1985), that there is little or no purely formal variation in language, directs the linguists attention to where it does not belong. Variation is simply part of the essence of language and linguistic evolution, and the fact that there seems to be a certain cognitive 'fear' among language users of fluctuations in what Von Humboldt referred to as 'äuere Sprachform', i.e. 'outer form', without corresponding functional or semantic differences,[21] does not in any way prove the contrary. Linguistic change which does not originate in borrowing presupposes quantitative variation between formally different (and structurally motivated) but otherwise entirely random options.

Notes

1. The original work on this paper was made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences. The study of the Rimburg dialect, which is central in this contribution, was supported by the former Foundation for Linguistic Research, which was funded by the Netherlands Organization for Scientific Research. Thanks are due to Geert Booij, Beth Hume and Marc van Oostendorp for discussion as well as to the anonymous reviewers and the editors for their very helpful comments. The usual disclaimers apply.
2. But see the considerations in the last paragraph of section 1.1 in Labov 2007.
3. Cf. Hinskens 1992: Ch. 3 for a more extensive account of the recent economic and socio-demographic history of the region as well as a sketch of the dialect situation.
4. See, however, Hinskens 1992: § 8.5.2.
5. In each case, the first form is a broad phonetic rendering of the dialect variant; this is followed by the orthographic representation of the standard variant and the gloss(es).
6. Except for /kɑn/, in the dialects involved no single 3. sg. pres. indic. inflected verb form has a rhyme consisting merely of a short vowel followed by /n/.
7. Cf. Hinskens 1992: § 8.4.2 for a more detailed discussion of this aspect. The variant of the preposition *aan* ('on') with a long vowel is invariably n-less. Cf. Hinskens 1992: § 8.4.1 for a full account of variable versus categorical n-lessness and vowel length in function words.

8. Analyses of dialect use elicited for this study -cf. section 6 below- brought to light that there is no significant difference in the relative amount of deletion between monomorphemic words and the first members of a compound word, witness the overall n-deletion ratios of 38.27 and 41.90, respectively, on a scale ranging between 0 and 100 (t=-.94 df=26 p=.358).
9. For German, see Wiese's (1986: 4) formulation.
10. The same holds for a cluster of dialects of German spoken along the Rhine river - including the Ripuarian dialects of German. Cf. Schmidt 1986.
11. Constisting of two or more prosodic words (cf. § 4.1 above), each with its original word stress.
12. This rule type, which refers directly to syntax, deals with the residue of cases that are not subject to prosodic conditioning. In this approach, "phonologically idiosyncratic items can be inserted into a phrasal context, subject to subcategorisation frames which freely make reference to phonological, morphological, and syntactic properties of the environment" (Crysmann, 2001: 3). These frames condition lexical phonological rules. Cf. also Kaisse's (1985) P1 postlexical rules. See Kaisse and Hargus 1993 for an overview of the main distinctions between lexical and postlexical rules.
13. The prosodic analysis of the two blocking conditions to be presented in this section differs from the one in Hinskens 1992: § 8.4.4 in that the Clitic Group (a level in the prosodic hierarchy developed in Nespor and Vogel 1986) is no longer included as a prosodic constituent. In this respect, the following analysis is based on Booij's (1996) empirically thoroughly grounded proposal to analyse Dutch proclitics as word-level adjuncts and enclitics as incorporated on the foot level. This choice is not a principled one, however, as there appear to be reasons to not dismiss the Clitic Group altogether; in this connection cf. Lloret 1997 (section 3.1) on cross-dialectal variation in shwa epenthesis in Catalan infinite verb forms. For didactic reasons, the representations of prosodic constituency in (12a), (17), (19) and (21) also contain grammatical categories such as V, Pro, P, Art, N and Dem.
14. A similar phenomenon, typical of a cluster of East-Limburg dialects, is central in § 4.2 of Wetzels 1993.
15. For brevity's sake the syllable nodes, which are not crucial at this point in the analysis, will not be included in the prosodic tree diagrams.
16. Either generally (Geert Booij, personal communication), or solely if they are stressed (cf. Neijt 1985: 185).
17. This representation integrates the above considerations regarding the partial feature identity of n] and [d on the one hand (also explicitly pointed out in Hinskens 1992: § 8.4.4), on the basis of Clements and Hume's 1995 model of feature organization, and the prosodic restrictions on n-deletion elaborated in this section on the other.

18. Cf. Harnisch's (1994) study of the relationship between frequency of usage on the one hand and deletion processes and stem allomorphy on the other, based on data for the German dialect of Ludwigsstadt. Bybee (2001) presents a model of the psycholinguistics of usage, including the effects of both type and token frequency.
19. A comparison with the results reported in Table 4 above shows that the part of the variance in n-deletion bound by the speakers' educational background is higher (before a consonant-initial word) or even much higher (in a word with a lexical High tone) than that bound by the speakers age groups.
20. To have more certainty in this connection, it would have to be clarified whether n-deletion, either generally or in specific linguistic conditions, shows stylistic variation.
21. Nicely illustrated in the 'vases', [vazɪz] ~ [vezɪz], example in Labov 1972: 251. For a parallel example see Hinskens 1992: § 1.2 4.

References

Auer, P. and F. Hinskens
1996 New and not so new developments in an old area, In: *The convergence and divergence of dialects in Europe*, P. Auer, F. Hinskens and K. Mattheier (eds.), *Sociolinguistica* 1996.

Auer, P., F. Hinskens and P. Kerswill (eds.)
2005 *Dialect change. Convergence and divergence in European languages.* Cambridge: CUP.

Booij, G.
1996 Cliticization as prosodic integration: The case of Dutch. *The linguistic review* 13: 291-242.

Booij, G. and D. de Jong
1987 The domain of liaison: theories and data. *Linguistics* 25: 1005-1025.

Chomsky, N. and M. Halle
1968 *The Sound Pattern of English.* New York.

Bybee, J.
2001 *Phonology and language use*. Cambridge: CUP.

Clements, G.N. and E. Hume
1995 The internal organization of speech sounds In J. Goldsmith (ed.) (1995), 245-306.

Coseriu, E.
1987 Über Leistung und Grenzen der Kontrastiven Grammmatik, In: *Formen und Funktionen. Studien zur Grammatik*, K. Baumgärtner (ed.), 67-84. Tübingen.

Coseriu, E.
1988a *Einführung in die Allgemeine Sprachwissenschaft.* Tübingen.

Coseriu, E.
1988b Interdisziplinarität und Sprache, In: *Energeia und Ergon. Sprachliche Variation - Sprachgeschichte - Sprachtypologie.* Band I. Schriften von Eugenio Coseriu (1965-1987). Eingeleitet und herausgegeben von J. Albrecht. Tübingen, 81-101.

Crysmann. B.
2001 Phonological properties of Portuguese clitics: a declarative approach. Paper presented at the 8th International Conference on HPSG, Aug 3-5 2001, Trondheim, Norway. http://www.coli.uni-saarland.de/~crysmann/papers/phonprop.pdf.

De Jong, E. (ed.)
1979 *Spreektaal; woordfrequenties in gesproken Nederlands.* Utrecht.

García, E.
1985 Shifting variation. *Lingua* 67(2-3): 189-224.

Gilles, P.
2006 Phonologie der n-Tilgung im Moselfränkischen (‘Eifler Regel’). Ein Beitrag zur dialektologischen Prosodieforschung. In *Perspektiven einer linguistischen Luxemburgistik. Studien zur Diachronie und Synchronie*, C. Moulin and D. Nübling (eds.), 29-68, Heidelberg: Winter.

Goldsmith, J. (ed.)
1995 *A handbook in phonological theory.* Oxford.

Guy, G.
1997 Competence, performance, and the generative grammar of variation, In: F. Hinskens, R. van Hout and L. Wetzels (eds.) (1996), 125-143.

Harnisch, R.
1994 Allomorphische Variation und kognitive Kosten. Evidenz aus dem Ortsdialekt von Ludwigsstadt in Oberfranken. In *Verhandlungen des Internationalen Dialektologenkongresses Bamberg 1990, Band 3*, W. Viereck (ed.), 283-96. Stuttgart.

Hayes, B.
1986 Inalterability in CV phonology. *Language* 62: 321-351.

Hayes, B.
1990 Precompiled phrasal phonology. In *The phonologysyntax connection*, Sh. Inkelas and D. Zec, (eds.), 85-108. Chicago.

Hermans, B.
1984 Het Limburgs en het Litouws als metrisch gebonden toontalen. *Spektator* 14(1): 48-70.

Hermans, B.
1994 *The Composite Nature of Accent. With Case studies of the Limburgian and Serbo-Croatian Pitch Accent*, Ph. D. thesis, Vrije Universiteit Amsterdam.

Hinskens, F.
1992 *Dialect levelling in Limburg. Structural and sociolinguistic aspects.* Ph. D. thesis, University of Nijmegen. An abridged and revised version was published under the same title by Niemeyer, Tübingen (1996).

Hinskens, F.
1998 Dialect levelling: a two-dimensional process. In *The role of the standard language in processes of convergence and divergence between dialects*, P. Auer (ed.), 35-51. Special issue of *Folia Linguistica*.

Hinskens, F.
2001 Koineization and creole genesis. Remarks on Jeff Siegel's contribution. In *Creolization and contact*, N. Smith and T. Veenstra (eds.), 199-218. Amsterdam/Philadelphia: Benjamins.

Hinskens, F.
in prep. Final n-deletion after lax vowels in monosyllabic words in Ripuarian and Limburg dialects of Dutch. Variability and blocking.

Hinskens, F., P. Auer, and P. Kerswill
2005 The study of dialect of convergence and divergence. Conceptual and methodological considerations. In P. Auer, F. Hinskens and P. Kerswill (eds.) (2005), 1-48.

Hinskens, F., R. van Hout and L. Wetzels (eds.)
1997 *Variation, change and phonological theory.* Amsterdam / Philadelphia.

Inkelas, Sh. and D. Zec
1995 Syntax-phonology interface. In J. Goldsmith (ed.) (1995), 535-549.

Itô, J.
1990 Prosodic minimality in Japanese. In *CLS 26, Part II. Papers from the Parasession on the syllable in phonetics and phonology*, K. Deaton et al. (eds.), 213-239. Chicago.

Kaisse, E.
1985 *Connected speech: the interaction of syntax and phonology.* New York

Kaisse, E. and Sh. Hargus
1993 Introduction. In *Studies in lexical phonology. Vol. 4 of Phonetics and phonology*, Sh. Hargus and E. Kaisse (eds.), 1-19. San Diego: Academic Press.

Kenstowicz, M.
1994 *Phonology in generative grammar.* Oxford.

Kiparsky, P.
1971 Historical linguistics, reprinted. In *Explanation in phonology*, 57-81. Dordrecht.

Kiparsky, P.
MS Uses and abuses of variation in historical explanation. Stanford University.

Labov, W.
1969 Contraction, deletion and inherent variability of the English copula. *Language* 45(4): 712-762.

Labov, W.
1972 *Sociolinguistic patterns.* Philadelphia.

Labov, W.
1991 The three dialects of English. In *New ways of analyzing sound change*, P. Eckert (ed.), 1-44. San Diego.

Labov, W.
2007 Transmission and diffusion. *Language* 83(2): 344-387.

Labov, W., S. Ash and C. Boberg
1995 A problem of macrosociolinguistics: uniformity of the North vs. diversity of the North Midland. Paper presented at the 24th Annual Conference on New Ways of Analyzing Variation (NWAVE), University of Pennsylvania, Philadelphia.

Labov, W., S. Ash and C. Boberg
2006 *The Atlas of North American English: Phonetics, phonology and sound change.* Berlin Mouton de Gruyter.

Lloret, M.-R.
1997 When does variability become relevant to linguistic theory? In F. Hinskens, R. van Hout and L. Wetzels (eds.).

Mattheier, K.
1980 *Pragmatik und Soziologie der Dialekte; Einführung in die kommunikative Dialektologie des Deutschen.* Heidelberg.

Murray, R. and Th. Vennemann
1983 Sound change and syllable structure in Germanic phonology. *Language* 59(3): 514-528.

Neijt, A.
1985 Clitics in arboreal phonology. In *Advances in nonlinear phonology*, H. van der Hulst and N. Smith (eds.), 179-192. Dordrecht.

Nespor, M. and I. Vogel
1986 *Prosodic phonology.* Dordrecht.

Rousseau, P. and D. Sankoff
1978 A solution to the problem of grouping speakers. In *Linguistic variation. Models and methods*, D. Sankoff (ed.), 97-117. New York.

Sapir, E.
1921 *Language.* New York.

Schmidt, J.
1986 *Die mittelfränkischen Tonakzente (Rheinische Akzentuierung).* Wiesbaden.

Selkirk, E.
1980 Prosodic domains in phonology: Sanskrit revisited. In *Juncture*, M. Aronoff and M.-L. Kean (eds.),107-129. Saratoga: Anma Libri.

Siegel, J.
1985 Koines and koineization. *Language in society* 14(3): 357-78.

Vallen, T. and Sj. Stijnen
1996 Terug naar Kerkrade. Een zoektocht naar sporen van het Kerkradeproject. In R. van Hout and J. Kruijsen (eds.) 239-59.

Van Hout, R.
1989 *De structuur van taalvariatie. Een sociolingustisch onderzoek naar het stadsdialect van Nijmegen*, Ph. D. thesis, KU Nijmegen.

Van Hout, R. and J. Kruijsen (eds.)
1996 *Taalvariaties. Toonzettingen en modulaties op een thema. Feestbundel ter gelegenheid van de zestigste verjaardag van Toon Hagen.* Dordrecht.

Van Oostendorp, M.
1995 *Vowel quality and phonological projection.* Ph. D. thesis, Tilburg University.

Van Oss, F. and C. Gussenhoven
1984 De Nederlandse slot-n in het nieuws. *Gramma* 8(1): 37-45 .

Vennemann, Th.
1988 *Preference Laws for Syllable Structure and the Explanation of Sound Change*. Berlin: Mouton de Gruyter.

Weijnen, A.
1991 *Vergelijkende klankleer van de Nederlandse dialecten.* ’s-Gravenhage.

Weinreich, U., W. Labov and M. Herzog
1968 Empirical foundations for a theory of language change. In *Directions for historical linguistics; a symposium*, W. Lehmann and Y. Malkiel (eds.), 95-189. Austin.

Wetzels, W.L.
1993 La phonologie de la flexion adnominale dans un dialecte Limbourgeois (Pays-Bas). In *L’architecture des représentations phonologiques*, B. Laks and A. Rialland (eds.), 203-231. Paris.

Wiese, R.
1986 Zur Theorie der Silbe. *Studium Linguistik* 20: 1-15.

Wolf, N.
1983 Durchführung und Verbreitung der zweiten Lautverschiebung in den deutschen Dialekten. In *Dialektologie. Ein Handbuch zur deutschen und allgemeinen Dialektforschung. Band 2*, W. Besch, U. Knoop, W. Putschke and H. Wiegand (eds.) 1116-1121. Berlin.

Wolfram, W. and N. Schilling Estes
1995 Moribund dialects and the endangerment canon: the case of the Ocracoke Brogue. *Language* 71(4): 696-721.

Wrede, F.
1919 Zur Entwicklungsgeschichte der deutschen Mundartforschung. *Zeitschrift für deutsche Mundarten* 20: 3-18.

Minimal morpheme expression in Dutch dialectology

Marc van Oostendorp

1. Introduction

One of the fundamental problems for constraint-based theories of phonology is the issue of opacity: a phonological process applies where it should not, or does not apply where it should, if we look at the phonological context on the surface. Since well-formedness constraints in a theory such as Optimality Theory typically refer to surface structure only, it is a puzzle how to deal with such phenomena within this theory. Various solutions have been proposed, all of them necessarily requiring some level of abstraction. Most of these involve some mechanism to introduce an extra representational level; an example of this is Sympathy Theory (McCarthy, 1999), in which one of the non-winning candidate representations can still influence the form of the output because it does contain the relevant phonological context, and the actual output form is bound to this candidate by faithfulness relations. This representation is abstract in the sense that it is not pronounced; but it can influence the shape of the surface form; the solution is highly reminiscent of derivational theories, in which the abstract representation is a separate step in the derivation.

Leaving aside the question whether this level of abstractness is required for other cases, I argue that one class of cases of phonological opacity can be handled without stipulating an extra level of representation, but by taking into account the morphological structure of the forms in question. In particular, deleted segments sometimes still seem to influence the surface representation of morphologically complex words, since without this influence a whole morpheme would be lost. I argue that there is a principle of the following general shape:

(1) *Phonological recoverability*. Every morpheme in the input should be represented in the phonological output.[1]

A functional explanation of (1) is possible, if needed: if a morphologically complex form needs to be parsed, it is preferable to have cues in the phonological shape for every independent morpheme, but (1) can also be seen as a purely formal requirement on linguistic structure, perhaps a consequence of some more general principle of the architecture of the language faculty. In particular, it can be seen as an instance of what Jackendoff (1993) calls 'correspondence rules' between components of grammar; Jackendoff makes it clear that such rules satisfy a conceptual necessity under any view of the grammar.

In this article, I propose to formalize Recoverability as an OT constraint:

(2) EXPRESS-[F]: The morphological feature F should be expressed in the phonological surface.
(Some phonological feature connected to the input expression of F should be present in the output.)

This is a special type of faithfulness constraint, basically stating that it is not allowed to delete a morheme fully. It could be formalised in various ways, depending on one's specific assumptions on the theory of faithfulness. van Oostendorp (2005) presents a formalisation in terms of the theory of Containment of Prince and Smolensky (1993); here, we will provide a definition in terms of Correspondence Theory (McCarthy and Prince, 1995a), as this is the more popular framework:

(3) EXPRESS-[F] (Correspondence Theoretic definition):
Given an input representation morpheme α and an output representation β.
If the morphological specification of α is F, and the phonological specication of α is a set of phonological features $\mathscr{F}$; then at least one feature $\phi \in \mathscr{F}$ should have a correspondent in β.

In other words, at least some part of the phonological specification of an underlying morpheme should survive to the output.[2]

Notice that an interface theory based on EXPRESS-[F] implies an Item-and-Arrangement view of morphology (Hockett, 1958): all morphemes are triples of sets of phonological, morphosyntactic and semantic features. EXPRESS-[F] constraints are special cases of MAX-F constraints, with an existential quantifier: they don't require *all* features of a morpheme to surface; it is enough if just one of them does.[3]

EXPRESS-[F] can be satisfied in many different ways, since it is not specified *which* feature needs to be present in the surface. The choice will be made by the independent ranking of markedness and faithfulness constraints in the grammar of the language: the feature which is least marked and/or for which we can establish the highest specific faithfulness requirement – even if this is ranked below EXPRESS-[F] – will suffice to fill the needs of this constraint. Similarly, EXPRESS-[F] are usually satisfied locally, i.e. close to the position where the original morpheme would have been expressed. I assume that the reason for this is the same ranking of constraints that normally function to ensure the order of morphemes, e.g. ALIGN-morpheme constraints (McCarthy and Prince, 1995b).

It can be shown how a number of apparent cases of phonological opacity can be dealt with if we use this mechanism. My examples in the following are taken from the literature on various Dutch dialects. These have been fairly well-studied in the Dutch dialectological literature, but are not well-known outside of this tradition.

I have made a further restriction to *inflectional* morphology. The reason for this is that inflection usually is rather 'weak' in the sense that inflectional morphemes in Germanic consist of only a few segments, and these are strongly susceptible to deletion: the vowel is usually schwa, the consonants are very often coronals, and rules deleting schwas and coronal consonants abound. If any morpheme ever is a likely candidate for violating (1), it most likely is an inflectional morpheme. Furthermore, we have a relatively clear view of the internal morphological structure of inflectional elements (which consist of purely 'formal' features only), whereas this is much less the case for derivational affixes, in which some amount of lexical semantics is also involved.

The structure of the argumentation will be the same in each example. An inflectional morpheme is phonologically weak in the way just outlined and therefore bound to be deleted. At the same time, if it would be present, it would either trigger or block a process of assimilation. In order to satisfy the requirement in (1) the deletion of the morpheme is not complete; the constituent of the original segment which is necessary to participate either positively or negatively in the assimilation process is left behind as a trace. For example, in Hellendoorn Dutch, an otherwise active process of progressive nasal consonant assimilation seems to be blocked in the past tense (in the cases below, the plural suffix may be assumed to be syllabic /n/; the orthographic examples represent Standard Dutch):

(4)
- a. *werken* ‘(to) work’ [wɛrkŋ̩]
- b. *werkten* ‘(we) worked’ [wɛrkn̩]
- c. *hopen* ‘(to) hope’ [hopm̩]
- d. *hoopten* ‘(we) hoped’ [hopn̩]

As can be seen from the orthography, and as will become evident if we study other instances in Hellendoorn Dutch, the imperfective suffix underlyingly contains at least a coronal obstruent /t/. We can now analyze this as a case of rule opacity: first we have an assimilation rule, and afterwards a rule of t-deletion, obscuring the original environment of assimilation. This approach can also be mimicked in e.g. Sympathy Theory, invoking faithfulness to a candidate output in which the /t/ is still present, and the nasal is assimilated to that /t/. There is no clear explanation under such an approach, why the rules are ordered in this way, or why the grammar needs to be faithul to this particular candidate output.

An alternative approach is to assume that /t/ is not deleted fully, but leaves behind a trace, in the form of the feature [coronal], which is then realized on the nasal consonant. The reason for this could be a general requirement that linguistic structure should be visible and expressed, i.e. the principle in (1). The consequences of this approach are explored in this article. The discussion will be embedded within Optimality Theory, currently the most popular theory of input-output mapping; but (1) is virtually theory-independent and its effects could be couched within other frameworks as well.

2. Opacity and Optimality

The most important work on rule opacity phenomena within generative phonology was done by Paul Kiparsky, in a number of papers that have been collected in Kiparsky (1982). In this work, Kiparsky eventually defines *opacity* in the following way:[4]

(5) *Opacity*

A rule A → B / C __ D is opaque to the extent that there are surface representations of the form

1. A in environment C __ D, or
2. B in environment other than C __ D

The converse of opacity is termed *transparency* by Kiparsky. It is important to note that Kiparsky's conception of opacity was clearly rooted in the generative framework of Chomsky and Halle (1968). The 'rule' in the definition in (5) is stated in purely segmental terms: A, B, C and D are all feature bundles, that is to say, phonological segments.

There is currently a rich literature on opacity within Optimality Theory. Approaches to this phenomenon can be divided in three, partly overlapping groups:

1. Approaches which use separate representations to store the 'hidden information', either as intermediate steps in a derivation (e.g. Stratal or Derivational OT of Kiparsky (2000); Rubach (2000); Kiparsky (To appear) or Candidate Chain Theory of McCarthy (2006)) or as somehow independently generated candidates (e.g. Sympathy Theory of (McCarthy, 1999, 2003b)).

2. Approaches which use only one representation, which is somehow 'enriched'. The current proposal will be of this type, as are e.g. Prince and Smolensky (1993)'s Containment Theory, Goldrick (2000)'s Turbidity Theory or Revithiadou (to appear); van Oostendorp (to appear)'s Coloured Containment.

3. Approaches which use constraints which work on (potentially) relatively shallow input and output representations; such constraints will typically combine aspects of faithfulness and markedness constraints. Examples are Comparative Markedness (McCarthy, 2003a) and Constraint Conjunction (Lubowicz, 2000).

It is not clear *a priori* which of these approaches is to be preferred. The first and the third approach have the merit of keeping individual representations relatively clear, but this seems to come at the cost of computability and learnability problems.

Representational approaches to opacity usually work only on certain subcases of opacity, for instance those involving deletion. However, this is not necessarily an argument against them, since they may invoke assumptions which seem independently needed, taking into account the literature on prosodic organization, autosegmental structure and the interaction with morphological and syntactic boundaries. One representation, which is enriched by independently necessary elements, may then do the work of two poorer representations.

In Antwerp Dutch, for instance, we have a process velarising a nasal consonant in coda position. We also have a process shortening a vowel before the resulting velar nasal (Taeldeman, 1982). Both processes can be seen at work in the following examples:

(6)

a. *grune* ‘green’ [ɣrynə] ∼ *gruun* ‘green’ [ɣryŋ][5]
b. *schoenen* ‘shoes’ [sxunə] ∼ *schoen* ‘shoe’ [sxuŋ]

Interestingly, the velarisation process only applies to words with underlyingly long vowels, and not to words which have short vowels already underlyingly:

(7)

a. *kin* ‘chin’ [kin] ∼ *tien* ‘ten’ [tiŋ]
b. *zon* ‘sun’ [zɔn] ∼ *zoon* ‘son’ [zoŋ]

In order to describe this, we could write the following rules (following Taeldeman, 1982):

(8)

a. n → ŋ / Vː__
b. Vː → V / __ ŋ

When applied to the different inputs in the right order, these rules will yield the correct results. While strictly speaking the rule in (8a) is opaque in the sense of Kiparsky’s definition in (5),[6] this is only so because it has a rather unnatural shape: it is very uncommon in languages of the world for velar nasals to only show up after long vowels. As a matter of fact, most variants of Dutch (and Germanic) allow the velar nasal to occur only after short vowels. This is even true for Antwerp Dutch, at the surface; rule (8b) is responsible for that. A much more natural rule would therefore be the one in (9):

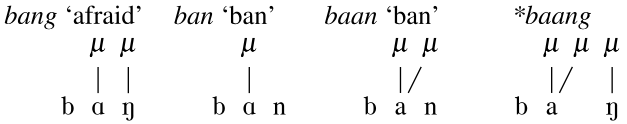
(9)

n → ŋ / V __

Yet this process would be very opaque indeed, since we obviously find cases where underlying /n/ did not turn into a velar.[7] We would therefore have to distinguish between underlyingly short and shortened vowels: the process seems so opaque that even an analysis based on rules (with arbitrarily many intermediate representations) cannot deal with it satisfactorily.

On closer inspection, there is evidence that the velar nasal, different from the other nasal consonants, is a moraic segment in many varieties of Dutch (Trommelen, 1982; van Oostendorp, 2001; van der Torre, 2003). This could explain, for one thing, the fact that velar nasals can only occur after short vowels: we could posit a bimoraic maximum on syllables.[8] A velar nasal after a long vowel would then be prohibited:

(10)



There undoubtedly are other ways to capture the same intuition: that both long vowels and velar nasals need space in the syllable and that if we would have both, this would be too much.[9] Under the one chosen here, the opacity of the Antwerp Dutch velarisation process disappears, if we assume that there is a strong faithfulness requirement on the number of mora’s in this dialect – or in rule-based terms, that we are not allowed to insert any moras in the underlying representation. A change from /tiːn/ – two underlying mora’s attached to the vowel – to [tiŋ] (two surface mora’s, one for the vowel and one for the nasal) is then allowed, but a change from /kin/ (one underlying mora attached to the vowel) to [kiŋ] (two surface mora’s) is not. If we allow ourselves to introduce a few ad hoc constraints to make things work technically, an OT analysis might then run along the following lines. We need the constraints in (11), and the ranking in (12) in order to get the tableau in (13):[10]

(11)

a. FAITH(μ): Do not add or delete mora’s.
b. VELAR: Nasal consonants in coda position should be ŋ.
c. *$\mu\mu\mu$: No trimoraic syllables.

(12) FAITH(μ)$\gg$VELAR

(13)

a.	☞	[tiŋ]			
		/tiːn/	Faith(μ)	$*\mu\mu\mu$	Velar
	i.	[tiːn]			*! W
	ii.	[tiːŋ]	*! W	*! W	
	iii.	[tin]	*! W		*! W

b.	☞	[kin]			*
		/kin/	Faith(μ)	$*\mu\mu\mu$	Velar
	i.	[kiːn]	*! W		*
	ii.	[kiːŋ]	*!* W	* W	L
	iii.	[kiŋ]	*! W		L

The winning candidate in (13a) does not violate any of the relevant constraints: it manages to keep the underlying mora structure (by moving the underlying second mora of the long vowel to the velar nasal), not to introduce a superheavy syllable and to end in a velar nasal at the same time. (13ai) is also fully faithful, but has a nonvelar nasal in coda; (13aii) adds a mora to the velar nasal which is not underlying and furthermore creates a superheavy syllable, whereas (13aiii) has deleted a mora from the underlying vowel and does not have a velar nasal.

The winning candidate in (13b) also does not feature a velar nasal, but in this case, there is no choice. Velarising the nasal would mean introducing a mora, as (13bii) and (13biii) show; lengthening the underlying vowel, as in (13bi) and (13bii) obviously does not improve the structure in any way.

All three constraints are known from the literature. No special mechanisms, such as rule ordering, Sympathy Theory, levels, etc., thus are necessary for Antwerp Dutch, given the appropriate assumptions on the representation of the velar nasal. This does not necessarily mean, of course, that all of the theoretical devices just mentioned are useless. There is an enormous literature on rule opacity, and many phenomena still defy explanation.

Yet in the following sections, I will show that next to a more sophisticated view of phonological structure, also a more precise view of the interaction between phonology and morphology, and in particular of the ‘visibility’ of morphology for phonology, may help to make many apparent examples of phonological opacity actually vanish. Constraints do not uniquely refer to segments, arranged in a one-dimensional string; what we actually have is a more fine-grained, multidimensional vision referring to the internal structure of these segments, and to the relations between them.

The analyses presented in this article are rooted in a tradition in which opacity is analysed as a form of recoverability: a process does or does not apply because otherwise the underlying shape of a form would be fully obscured. Representatives of this line of thinking are Kaye (1974); Gussman (1976); Kissebert (1976). Different from these scholars, we concentrate on a specific type of recoverability, viz. morphological recoverability: the application of phonological processes is dependent on the question whether or not all morphemes will be visibly present at the surface.

3. Nasal assimilation in Hellendoorn Dutch past tense

As outlined above, Hellendoorn Dutch – a dialect spoken in the northeastern parts of the Netherlands –, like many other languages in the world, displays a process of nasal assimilation. Interestingly, the process works from right to left as well as from left to right. The following facts are all from Nijen Twilhaar (1990), the orthography again is Standard Dutch:

(14)

	orthography	underlying	surface	gloss
a.	*lopen*	lopən	lopm̩	‘to walk’
b.	*weten*	wetən	wetn̩	‘to know’
c.	*pakken*	pɑkən	pɑkŋ̩	‘to grab’
d.	*loop een*	lop ən	lopm̩	‘(I) walk a (mile)’
e.	*rampnacht*	rɑmpnɑxt	rɑmpnɑxt	‘disastrous night’
f.	*loop een keer*	lop ən ker	lopŋ̩kɪːr	‘(I) walk one time’

Nasal assimilation in Hellendoorn Dutch has some interesting properties. Examples (14a-c) show that a (syllabic) nasal assimilates to a preceding obstruent. In contradistinction to the first analysis in (4), the plural suffix is represented here as underlyingly /ən/. I will return to this assumption below. For now it suffices to see (14d) that the indefinite article, which unquestionably has a schwa underlyingly (because this schwa surfaces e.g. if an indefinite nominal phrase occurs at the beginning of a sentence), displays the same behaviour. (14e) shows that nasals in onset position are not affected by the process, and (14f) that in certain cases assimilation is *regressive*, to a following consonant.

The key facts are the ones in (15):

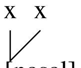
(15)

a. /stɔp+t+n/ (stop+past+plural) >[stɔpn̩] ‘stopped’
b. /zɛt+t+n/ (put+past+plural) >[zɛtn̩] ‘put’
c. /pɑk+t+n/ (grab+past+plural) >[pɑkn̩] ‘grabbed’

This is a case of opacity because within a rule-based framework, we could state two rules (disregarding regressive assimilation), one of progressive nasal assimilation, and another one of /t/ deletion (the following is based on Nijen Twilhaar, 1990):

(16)

a. t deletion: t → 0 / C __ C



b. progressive assimilation (PA): [nasal]

PA is rendered opaque by t deletion (schwa deletion is implied to proceed the processes described here):[11]

(17)

	/stɔp+t+n/	/zɛt+t+n/	/pɑk+t+n/
schwa deletion	stɔptn̩	zɛttn̩	pɑktn̩
PA	stɔptn̩	zɛttn̩	pɑktn̩
t deletion	stɔpn̩	zɛtn̩	pɑkn̩

It is fairly easy to set up an analysis of the non-opaque facts in (14). Again, we use a few constraints which may not be hallmarks of theoretical sophistication, but which give the required results.[12]

(18)

a. ASSIMILATE: A coda nasal and an adjacent obstruent should have the same place of articulation.
b. *CCC: Clusters of three consonants are not allowed.[13]
c. FAITH(PLACE): Input place features of the stem should surface.

(19)

☞ pɑkŋ̩			
/pɑk+n/	*CCC	ASSIMILATE	FAITH(PLACE)
pɑkn̩		*! W	
pɑtn̩			*! W

Hellendoorn differs from other languages displaying faithfulness of place features in that even after the consonant deletion, another obstruent stays present that could still enforce assimilation. Therefore, the opaque cases here cannot be dealt with without additional means:[14]

(20)

pakn̩		*!	
/pak+t+n/	*CCC	Assimilate	Faith(Place)
✗ pakŋ̍		L	
paktn̩	*! W		*W

What we need to express, here, is the idea that the nasal gets its feature from the underlying past tense suffix. In order to achieve this, we introduce a constraint on the expression of the morphological feature [tense]:

(21) Express-[Tense]: The morphological feature tense should be expressed in the phonological surface.

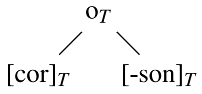
As we argued in the introduction, this is a special type of faithfulness constraint: some part of the past tense suffix /-t/ should be expressed in the output. Adding this constraint to our tableau gives us the desired result:

(22)

☞ pakn̩			*	*
/pak+t+n/	*CCC	Express-Tense	Assimalte	Faith(Place)
pakŋ̍		* W	L	*
paktn̩	* W		L	L

How exactly does Express-Tense work? The phonological input is a /t/, i.e. a feature bundle like the following:

(23)



Every morpheme consists of a number of feature bundles, connected to a root note, and/or a timing slot. I marked this by adding a subscript to every individual element. Seen in this notation, Express-Tense states that the output should contain at least one element which has this subscript.[15]

Until this point, we have silently assumed that the plural suffix is /n/. However, it also seems possible to assume /ən/ has a schwa underlyingly. We can find some arguments for this in Nijen Twilhaar (1990). Most convincing perhaps is the argument that we also find monomorphemic nouns ending in a syllabic nasal, and nouns ending in /-ə/, but no nouns ending in /-ən/. This shows at least that a productive process of schwa deletion before n is going on. Furthermore, the schwa sometimes surfaces, viz. in very formal styles of speech (Nijen Twilhaar, 1990: 165); these are styles where typically the surface form is closer to the underlying representation (van Oostendorp, 1997).

There are various reasons why schwa should not surface; being phonologically and phonetically empty, it seems a less desirable nucleus, etc. There will thus be a constraint *ə or something more motivated but to the same effect (cf. van Oostendorp, 2000, for fuller detail).

We can distinguish betwee three groups of dialects of Dutch (van Hout and van der Velde, 2000). In some, schwa is deleted, and in others /n/ is deleted under various circumstances; in line with the previous discusssion, this could be formalized as a constraint *n. The third variety is one in which neither schwa nor /n/ is deleted. Crucially absent are those dialects in which *both* schwa *and* /n/ are deleted. This may be seen as an indication for the high level of activity of recoverability, formalized in this case as a constraint Express-Plural. This constraint would then dominate at least one of *ə and *n.

These facts are not unique for Hellendoorn Dutch; we find very similar phenomena even in typologically unrelated languages. The interaction between nasal assimilation and consonant deletion from the Ojibwa dialect of Odawa (Piggott, 1974; Kaye, 1974) show a very similar pattern:

(24)

Underlying	/takossin-k/ ‘he arrives’
Assimilation	takoʃʃiŋk
Deletion	takoʃʃiŋ

These facts are clearly very similar to those of Hellendoorn in the relevant respects. Viewed from a purely segmental point of view, Assimilation and Deletion are in an opaque (‘counterbleeding’) order. Yet if we consider the possibility that the place feature [velar] is on an independent plane, and that it can be retained even after deletion of the segment /k/, these facts follow.

Also in this case, it appears that /k/ is an independent morpheme, having a conditional meaning. If we assume that this conditional morpheme has to be retained at the surface somehow, the feature [velar] would then show up on the nasal as a trace of this morpheme in order to satisfy Express-[cond]. All in all, we would get an analysis such as the following:

(25)

a. Express-[cond]: The conditional morpheme should be expressed at the surface

b. *CC: Consonant clusters are not allowed.[16]

The tableaux of Hellendoorn and Ojibwa thus become strikingly similar. Again, there is no need to rank most of the relevant constraints in order to get the difference between conditionalis and realis forms:

(26)

☞ takoʃʃin				
/takossin/	*CC	Express-Cond	Assimilate	Faith(Place)
takoʃʃiŋ				* W

(27)

☞ takoʃʃiŋ				*
/takossin+k/	*CC	Express-Cond	Assimilate	Faith(Place)
takoʃʃin		*! W		L
takoʃʃink	*! W		* W	L
takoʃʃiŋk	*! W			*

As a matter of fact, there thus is no opacity, or any problem for surface-based phonology within Hellendoorn Dutch or Ojibwa at all, given the fairly standard assumption that place or articulation features can exist independently of their segments – an assumption that was not available in the work of Kiparsky and Kaye just referred to above.

4. Voicing Assimilation in Flanders and Brabant

The next example which deserves discussion is widespread in the Dutchspeaking parts of Belgium (at least in Flanders and Brabant). In this case a process of voicing assimilation interacts with the deletion of word final /t/, which is the phonological shape of the third person singular verbal inflection (Taeldeman, 1982) in a way that may be considered opaque:[17]

(28)

a. [-sonorant, +continuant] → [-voice] / [-sonorant] __

b. $t_{3,sg} \rightarrow \emptyset$ __ # C

(29) *hij doe/t v/eel* 'he does a lot'

a. *hij doe/t f/eel*

b. *hij doe[f]eel*

(28a) is an instance of Kiparsky's case (ii): we find [-voice] in an environment where we would not expect it. In derivational terms, it looks as if we first have assimilated the fricative to the preceding voiceless plosive, and then deleted the plosive. But this obviously means that the source of voicelessness is no longer visibly present on the surface.

Voicing assimilation in Dutch is a well-known and widespread phenomenon (cf. Lombardi, 1999; van der Torre and van de Weijer, in press, for various analyses within the OT framework). Lombardi (1999: 277) analyses it in terms of the following constraint:

(30) AGREE: Obstruent clusters should agree in voicing.

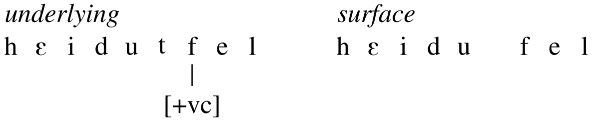
AGREE is of course very similar in form and spirit to the constraint ASSIMILATE, which we used above to describe nasal assimilation. Under this formulation, the process in question becomes unmistakenbly opaque, quite independent where we rank AGREE(if *CC≫FAITH(VOICE)):

(31)

[f]			*!
*hij doe/*t v*/eel*	*CC	AGREE	FAITH(VOICE)
[tv]	*! W	*W	L
[tf]	*! W		*
✘ [v]			L

In order for a solution along the lines of EXPRESS-[F] (in this case: EXPRESS-[3PS]) to work, we need to know what exactly is the phonological element that expresses the inflectional suffix in this case. Lombardi (1999) employs a theory of laryngeal features in which [voice] is monovalent, i.e. there is no phonologically active [-voice]. This means that the underlying and surface representations are schematically as follows (representing [v] as a /f/ with attached [voice]):

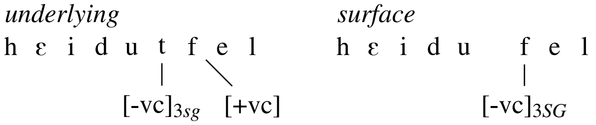
(32)



The underlying representation of the [3sg] suffix seems to have disappeared without leaving a trace; there is nothing in the surface form to represent it, given the plausible assumption that absence of a feature cannot act as a representative.[18]

Wetzels and Mascaró (2001) and various other authors have argued on independent grounds that there are empirical arguments to assume that [-voice] should be assumed to be phonologically present. In that case, the inflectional suffix does indeed leave a trace at the surface representation, viz. the feature [-voice], realized on the [f]:

(33)



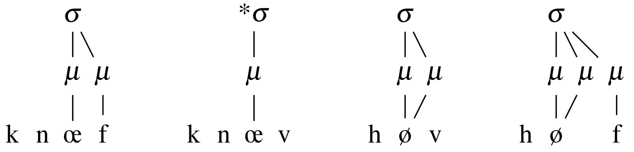
In van Oostendorp (2003), on the other hand, it is argued that the difference between voiced and voiceless fricatives phonologically behaves like a length distinction in many Westgermanic dialects. Intervocalically, at least, 'voiced' fricatives are short, and 'voiceless' fricatives are long. This explains, among other things, why in many of these dialects we find voiced fricatives after tense or long vowels and voiceless fricatives after lax or short vowels (the following examples are from Dutch):[19]

(34)

knuffel 'hug' [knœfəl]	*[knœvəl]
heuvel 'hill' [høvəl]	*[høfəl]

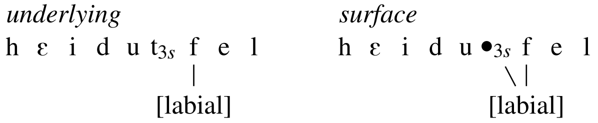
These facts can be accounted for straightforwardly if we assume that (stressed) syllables are minimally *and* maximally bimoraic, that long vowels occupy two mora's, and that short vowels and long consonants each occupy one mora (the structure of the syllable headed by schwa will be left out of consideration here):

(35)



If this analysis is correct, it could be the *position* of the /t/ that is retained after the disappearance of the segment:

(36)



It thus is the phonotactic position (represented here as a dot, since it does not matter whether this is a mora, a root node, or something else), that could be seen as the trace of the suffix, necessary to satisfy Express-3sg.

(37)

☞ [f]				*
*hij doel*t v*leel*	*CC	Express-3sg	Agree	Faith(Voice)
[tf]	*! W			*
[tv]	*! W		* W	L
[v]		*! W		L

Notice that the tableau is very similar to the ones given for Hellendoorn and Ojibwa above: the morphological faithfulness constraint Express-3sg outranks the phonological faithfulness constraint Faith-voice.

Taeldeman (1982) asserts that the opacity effect attested in the dialects just discussed is typical for deleted /t/ as a morpheme. Similar effects can however be found in the literature purely internal to the phonology in other dialects of Dutch. E.g. in Wilsum Dutch, the final /t/ of function words such as *net* (just), *met* (with) and *det* (that) can disappear, but still have the effect of devoicing the following /f/. We thus get examples such as the following (Spa, 2000: 46):

(38)

a. *ne/t z/o dudelijk >ne[s]o dudelijk* ‘just as clear’
b. *me/t z/ien ome >me[s]ien ome* ‘with his uncle’
c. *as ie de/t v/raogen >as ie de[f]raogen* ‘if he that asks=if he asks that’

In this case, the relevant type of faithfulness cannot be morphological, since the function words in question are expressed by an onset consonant and a vowel; Express constraints are thus satisfied in any case. It could be argued that we thus have an instance of faithfulness to the segmental position, usually formalized as a Dep constraint (every segment in the input should have a correspondent in the output). The correspondent of *t* will be filled by material from the fricative (which means that a lower-ranking Identity-constraint will be violated).

(39)

☞ [s]				*
ne/t z/o dudelijk	*CC	Dep	Agree	Faith(Voice)
[ts]	*! W			*
[tz]	*! W		* W	L
[z]		*! W		L

The situation is somewhat more complicated, because we need to explain why this only involves the final /t/ of function words. It has been argued that the final /t/ of words such as these is also in some abstract sense a suffix (Vanacker, 1949); the /t/ shows other types of deviant behaviour in other contexts as well. Taeldeman and Schutter (1986: 114) propose that there is a hierarchy of positions where devoicing after deleted /t/’s may occur:

1. verbal inflection
2. the ‘small words’ *daT* ‘that’, *waT* ‘what’, *nieT* ‘not’ and (sometimes) *meT* ‘with’
3. frequent adjectives of the type ʊːd (e.g. *goed* ‘good’, *kwaad* ‘angry/bad’, *koud* ‘cold’, *dood* ‘dead’)

This hierarchy seems to reflect the straightforwardedness of the inflectional nature of the /t/ in question. The hierarchy is based on the fact that in some dialects (the ones just discussed) we only find this effect in environments (i), in other dialects (the one spoken in Bruges) we find them in (i) and (ii), and yet in others (e.g. Ghent) we find them in (i), (ii) and (iii).

(40)

	Bruges	Ghent
hij doe/t v/eel 'he does a lot'	[f]	[f]
da/t v/uur 'that fire'	[f]	[f]
goe/d z/aad 'good seeds'	[z]	[s]

The question then remains why only frequent adjectives of a specific shape participate in Ghent. The frequency effect might be attributable to the fact that there tends to be more deletion in frequent words than in less frequent words in general (Goeman, 1999). It is not exactly clear to me why there should be a preference for adjectives ending in long vowels plus /d/. Yet one aspect is of particular importance here: the fact that in this case an analysis in which an underlying feature [-voice] would be the trace of the deleted segment cannot work. This gives indirect support for the analysis presented above in which it is the position of the coronal stop that surfaces, filled with the material of the fricative, which thereby lengthens and thus remains voiceless.

5. Conclusion

In this article, I have argued that a class of opacity problems disappear once we assume a mildly enriched theory of phonological representations: the enrichment consists precisely and exclusively of noting the morphological affiliation of every phonological feature. If we 'know' for every segment and feature to which morpheme it belongs – or whether it is epenthetic, i.e. it does not belong to any morpheme at all – it is sufficient to have one representation.

In order to solve the problem of opacity, many Optimality Theoretic proposals have suggested to introduce extra levels of representation. In this way, the 'rule-based' solution to this problem (where the extra representations are seen as stages in the derivation) was essentially mimicked. In most cases, the analysts tried to relate the representation to already existing ones, such as other candidates in the tableau in the case of Sympathy. In spite of this, these approaches run against the spirit of monostratalism, in which all constraints should refer exclusively to one level of representation.

The solution of Recoverability is monostratal, in the sense that only the output of the Generator function is relevant. The cost of this is that this single output representation is in some sense 'enriched': it contains information as to the morphological structure of the word, in particular about which phonological segment belongs to which morpheme.

This does not mean, of course, that we can do away with all independent representations – and thus with Sympathy Theory, Paradigm Uniformity, etc. – immediately. Many phenomena have been analysed using these devices, whereas the representational enrichment presented here mainly serves to deal with cases where at least one of the two processes involves ‘deletion’ of material. It is not excluded that the output representations could be enriched sufficiently so that they would be able to deal with these phenomena as well, but this remains an empirical issue in need of further exploration.

Notes

1. Constraints which are similar to this in one way or another have been proposed among others by Samek-Lodovici (1993); Akinlabi (1996); Gnanadesikan (1997); Rose (1997); Walker (1998, 2000); Piggott (2000); cf. Kurisu (2001) for an overview.
2. Other constraints referring to the morphological affiliation of phonological material can also be found. A well-known example of this are the positional faithfulness constraints referring to stems and affixes (FAITH(AFFIX) and FAITH(STEM)). In order to be able to check whether a changed or deleted segment violates either of these constraints, we need to know whether the segment belongs to a stem of to an affix. We need something like the subscripts used here for recoverability. To give a more concrete example, Bakovic (2000) proposes the principle *Stem identity* in order to deal with such phenomena:

   (i) STEM IDENTITY. Given [stem$_i$] and [[stem$_j$]afx], where afx is a prefix or suffix and stem$_i$ is morphologically identical to stem$_j$, application of a process P across the stem$_j$ and afx boundary may not thereby render stem$_i$ and stem$_j$ phonologically distinct.

   Bakovic (2000) uses this constraint to explain stem-controlled vowel harmony (where the harmonic features spread ‘inside out’ in morphologically complex forms), and why stems are never affected by affix vowels in such systems. Compare e.g. the following Turkish forms:

   (ii) 1. Input: *ip* ‘rope’+*lar* ‘nominative plural’
        2. Outputs: *iplar, ipler, ıplar*

   (In reality, it is hard to decide whether *-ler* or *-lar* is the underlying form of the suffix; but the argument does not change if we assume that it should be *ler* and consider combinations with back vowels.)
   The candidate *iplar* is disharmonic and therefore unacceptable. The candidates *ipler* and *ıplar* on the other hand, are harmonic; the first one by virtue of the

suffix adapting to the root, the second one by the root adapting itself to the suffix. Stem identity now chooses for the former, since this preserves the phonological structure of the stem *ip*. In order for STEM IDENTITY to work, the phonological constraints need to be able to see that the vowel /i/ɪ/ belongs to the stem, whereas /e/a/ belongs to an affix. This is exactly the type of information that is needed for a constraint such as EXPRESS-MORPHEME as well.

3. Struijke (2000) also proposed existential quantification within faithfulness constraints, but in that case the existential quantifier is in a different place: her constraints say that *for every feature* in the input, there should be some correspondent *somewhere* in the output.
4. There is a third clause to the definition of opacity, 'instances of B not derived by - that occur in the environment C__D', which is often ignored in the literature, and will also be ignored here, since it does not seem to refer to what people usually consider to be opacity.
5. Apart from some cases of optional schwa deletion, the form without schwa can only be used in the singular neuter of adjectives in indefinite noun phrases. The form with schwa can be used in all other inflections (plural or non-neuter, or both).
6. Rule (8a) is opaque, because we find velarised consonants which are not in the environment of a long vowel. In other words, clause (ii) is violated: we find B (=ŋ) in a context different than C__D (where C=Vː and D=Ø).
   Rule (8b) is strictly speaking also opaque according to clause (ii), since we find B (is a short V) in a context different from C__D, where C=Ø and D=ŋ (for instance, we also find short vowels before a coronal). This only applies to this strict definition, it disappears if we consider that all short vowels which are *the result of shortening* (hence underlying long vowels) occur only in this particular context.
7. In other words, the rule in (9) is opaque under clause (i) of Kiparsky's definition: we find an A (=n) in the context C__D, where C=V (a short vowel) and D=Ø.
8. It is sometimes assumed that the reason for this restriction is that velar nasals underlyingly are /nɡ/ or /nɣ/, just like this is assumed for English. The Antwerp facts actually show that this analysis cannot be correct: here we find the same restriction but there is absolutely no possibility for postulating an underlying obstruent.
9. Some approaches to Dutch syllable structure, such as Kager (1999), assume that all Dutch syllables have a bimoraic minimum; the form *ban* would have a moraic /n/. We could then assume that the difference between the velar nasal and the other two nasals is that the former, but not the latter, needs to have a mora 'of its own'. The /n/ would be able to project a mora if this is necessary, e.g. in [bɑn], but it would not need to do this in other cases, e.g. in [ban]. A bimoraic maximum would therefore still be responsible for the different distribution of the two segments.

Yet another possibility might be to include a constraint in our grammar which would have it that velar nasals have to occur in a nuclear position, leaving only one such position for the vowel (van der Torre, 2003).

10. I use the tableau format proposed by Prince (2002), since it seems more informative than the older tableau format. Here is how Prince explains it: "the winner and its violations appear above the constraints; the losing candidates are below. In addition to the losers' violations, the tableau shows each loser's relationship to the winner: does a constraint favor the winner over this loser (W), or does a constraint favor this loser over the winner (L), or does a constraint favor neither (blank). If a tableau is to yield the correct result, every loser row must contain at least one W with no L to its left. Since every loser-favoring constraint must be dominated by some winner-favoring constraint, ranking arguments involve checking that every L has a W to its left."
    It is the latter point that makes this tableau format very convenient: all crucial rankings can be read off tables which are presented in this format immediately. Cf. Prince (2002) for more arguments.
11. Interestingly, this is a case of opacity either of type (i) or of type (ii) in terms of Kiparsky's definition in (5), depending on how we look at it. We have surfacing CAD in the sense that it looks as if the nasal has not been subject to assimilation even though the context is present; we have XBY in the sense that it looks as if the nasal has been subject to assimilation to a segment which is no longer there. I have chosen the second possibility here. There would be ways to test which of these two theories is correct, if we would be able to find e.g. cases where the deleted consonant is non-coronal.
12. It would certainly be possible to give more sophisticated analyses using more elegant constraints, but these would require more different constraints, and the point would remain the same: an extra faithfulness constraint is necessary to understand the exceptional behaviour of past tense forms.
13. *CCC is a cover constraint for a variety of phonotactic constraints disallowing clusters of consonants, introduced here in order to simplify our tableaux. The constraint is assumed to be blind to both morphological and syllable boundaries.
14. It is assumed here that deletion of /k/ or /n/ is not an option for satisfying *CCC. As we will see below, /t/ is particularly prone to deletion in dialects of Dutch, but other consonants cannot be deleted. Exploring the reasons for this is beyond the scope of this article.
15. Most current correspondence-based theories of faithfulness are segment-based, i.e. the faithfulness to features is mediated through the segmental node. (There are no constraints MAX(CORONAL) or DEP(CORONAL), but only IDENT(CORONAL).) In order to implement the ideas presented here in such a framework, one would need a device like 'existential identity' (Struijke, 2000): "If an input segment x is [αF], then some output correspondent is also [αF]."

16. Like *CCC, this constraint is a cover constraint for a ‘real’ phonotactic analysis against clusters of consonants irrespective of syllable boundaries and morphological structure.
17. The rules are slightly simplified versions of the ones given in Taeldeman (1982). Taeldeman uses a definition of opacity which is somewhat different from the one employed here; under that definition, this rule ordering is actually not opaque, but the problem for Optimality Theory obviously stays the same.
18. The analysis presented here differs on this point from the one in Kurisu (2001), where the constraint REALIZEMORPHEME can be satisfied if the phonological shape of a derived form *differs* from that of the phonological shape of the base. Kurisu (2001) thus allows subtractive morphology, whereas the approach presented here does not. EXPRESS-[F] can only be satisfied if there is something in the underlying form which survives to the surface. Deletion is a change in terms of Kurisu (2001), but it does not satisfy EXPRESS-[F].
    This obviously makes an account in terms of EXPRESS-[F] more restrictive, although it remains to be seen, as always, whether hardcore cases of subtractive morphology could also be dealt with in this way.
19. There are a few apparent counterexamples to these generalisations, on which see van Oostendorp (2003). There is a debate in the Dutch literature as to the phonological representation of the relevant contrast in vowels: whether it is a feature such as [tense] or structural length. We adopt the latter position here (van Oostendorp, 2000). Some dialects have both a length contrast and a tenseness contrast; but those dialects are irrelevant to our present discussion.

References

Akinlabi, Akinbiyi
1996 Featural Affixation. *Journal of Linguistics* 32: 239–289.

Bakovic, Eric
2000 *Harmony, Dominance and Control*. Ph.D. thesis, Rutgers.

Chomsky, Noam and Morris Halle
1968 *The Sound Pattern of English*. Cambridge, MA: The MIT Press.

Gnanadesikan, Amalia
1997 *Phonology with Ternary Scales*. Ph.D. thesis, University of Massachusetts at Amherst.

Goeman, Ton
1999 *T-deletie in Nederlandse Dialecten. Kwantitatieve analyse van structurele, ruimtelijke en temporele variatie*. Ph.D. thesis, Vrije Universiteit Amsterdam.
http://www.meertens.nl/uitgeverij/extern/tonpdf/download.html.

Goldrick, Matthew

2000 Turbid Output Representations and the Unity of Opacity. *Proceedings of NELS*, 30: 231–245. Also available as ROA 368–1299.

Gussman, Edmund

1976 Recoverable derivations and phonological change. *Lingua* 40: 281–303.

Hockett, Charles

1958 Two models of grammatical description. In *Readings in linguistics*, Martin Joos (ed.) Chicago. University of Chicago Press.

van Hout, Roeland and Hans van der Velde

2000 N-deletion in reading style. In *Linguistics in the Netherlands 2000*, 209–219. Amsterdam: John Benjamins.

Jackendoff, Ray

1993 *The Architecture of the Language Faculty*. Cambridge, Mass.: The MIT Press.

Kager, René

1999 *Optimality Theory*. Cambridge: Cambridge University Press.

Kaye, Jonathan

1974 'Opacity and Recoverability in Phonology'. *Canadian Journal of Linguistics* 19: 134–149.

Kiparsky, Paul

1982 *Explanation in Phonology*. No. 4 in Publications in Language Sciences. Dordrecht: Foris.

Kiparsky, Paul

2000 Opacity and Cyclicity. *The Linguistic Review* 17: 351–367.

Kiparsky, Paul

To appear *Paradigmatic Effects*. Stanford: CSLI Publications.

Kissebert, Charles

1976 The Interaction of Phonological Rules and the Polarity of Language. In *The Application and Ordering of Phonological Rules*, Andreas Koutsoudas (ed.), 41-54. The Hague: Mouton.

Kurisu, Kazutaka

2001 *The Phonology of Morpheme Realization*. Ph.D. thesis, University of California at Santa Cruz.

Lombardi, Linda

1999 Positional Faithfulness and Voicing Assimilation in Optimality Theory. *Natural Language and Linguistic Theory* 17: 267–302.

Lubowicz, Anna

2000 Derived Environment Effects in Optimality Theory. *Lingua* 112: 243–280.

McCarthy, John
1999 Sympathy and Phonological Opacity. *Phonology* 16: 331–399. Also available as ROA 252-0398.

McCarthy, John
2003a Comparative Markedness. *Theoretical Linguistics* 29: 1–51.

McCarthy, John
2003b Sympathy, Cumulativity, and the Duke-of-York Gambit. In *The Optimal Syllable*, Caroline Féry and Ruben van de Vijver (eds.), 23–76. Cambridge: Cambridge University Press.

McCarthy, John
2006 *Hidden generalizations. Phonological opacity in Optimality Theory*. Los Angeles: Equinox.

McCarthy, John and Alan Prince
1995a Faithfulness and Reduplicative Identity. *UMOP* 18: 249–384.

McCarthy, John and Alan Prince
1995b Generalized Alignment. *Morphology Yearbook*.

Nijen Twilhaar, Jan
1990 *Generatieve fonologie en de studie van Oostnederlandse dialecten*. Amsterdam: P.J. Meertens-Instituut.

van Oostendorp, Marc
1997 Style levels in conflict resolution. In *Variation, Change and Phonological Theory*, Frans Hinskens, Roeland van Hout and Leo Wetzels (eds.), 207–229. Amsterdam: John Benjamins.

van Oostendorp, Marc
2000 *Phonological Projection*. Berlin/New York: Mouton de Gruyter.

van Oostendorp, Marc
2001 Nasal Consonants in Variants of Dutch and Some Related Systems. *Neerlandistiek.nl.* http://www.neerlandistiek.nl/01/08/index.html.

van Oostendorp, Marc
2003 Ambisyllabicity and Fricative Voicing in West-Germanic Dialects. In *The Syllable in Optimality Theory*, Caroline Féry and Ruben van de Vijver (eds.), 304–337 Cambridge: Cambridge University Press.

van Oostendorp, Marc
2005 The Theory of Faithfulness. Unpublished manuscript.

van Oostendorp, Marc
to appear Derived Environment Effects and Consistency of Exponence. In *Freedom of Analysis?*, Sylvia Blaho, Patrik Bye and Martin Krämer (eds.).

Piggott, Glyne
1974 *Aspects of Odawa Morphophonemics*. Ph.D. thesis, University of Toronto.

Piggott, Glyne
2000 Against Featural Alignment. *Journal of Linguistics* 36: 85–129.

Prince, Alan
2002 Arguing Optimality. http://roa.rutgers.edu/view.php3?roa=562. Also available as ROA 562-1102.

Prince, Alan and Paul Smolensky
1993 Optimality Theory: Constraint Interaction in Generative Grammar. Manuscript, Rutgers University and University of Colorado at Boulder. Available at ROA and published 2004 at Blackwell.

Revithiadou, Anthi
to appear Colored Turbid accents and Containment: A case study from lexical stress. In*Freedom of Analysis?* Sylvia Blaho, Patrik Bye, and Martin Krämer (eds.), Berlin: Mouton De Gruyter.

Rose, Sharon
1997 *Theoretical Issues in Comparative Ethio-Semitic Phonology and Morphology*. Ph.D. thesis, McGill University, Montreal.

Rubach, Jerzy
2000 Glide and glottal stop insertion in Slavic languages: A DOT analysis. *Linguistic Inquiry* 31: 271–317.

Samek-Lodovici, Vieri
1993 Morphological Gemination. Paper presented at Rutgers Optimality Workshop I.

Spa, J.J.
2000 *De dialecten van Groot-IJsselmuiden. Klank- en vormleer*. Kampen: IJsselacademie.

Struijke, Caro
2000 *Reduplication, Feature Displacement, and Existential Faithfulness*. Ph.D. thesis, University of Maryland.

Taeldeman, Johan
1982 Regelordening, taalverandering en ruimtelijke taalvariatie. *Taal en Tongval* 34(3/4): 120–131.

Taeldeman, Johan and George De Schutter
1986 Assimilatie van stem in de zuidelijke Nederlandse dialekten. In *Vruchten van zijn akker. Opstellen van (oud)medewerkers en oudstudenten voor Prof. V.F. Vanacker*, Magda De Vos (ed.), Gent: Universiteit Gent. http://www.dbnl.org/tekst/schu031assi01/.

van der Torre, Erik Jan
2003 *Dutch Sonorants: The Role of Place of Articulation in Phonotactics*. Ph.D. thesis, Leiden University.

van der Torre, Erik Jan and Jeroen van de Weijer (eds.)
in press *Voicing in Dutch*. Amsterdam: John Benjamins.

Trommelen, Mieke
1982 *The Syllable in Dutch: With Special Reference to Diminutive Formation*. Ph.D. thesis, University of Utrecht. Published at Dordrect: Foris, 1984.

Vanacker, V.F.
1949 Over enkele meervoudsvormen van voegwoorden. *Taal en Tongval* 1: 32–45, 77–93, 108–112.

Walker, Rachel
1998 *Nasalization, Neutralization, and Opacity Effects*. Ph.D. thesis, University of California, Santa Cruz. Published from Garland Press, 2000.

Walker, Rachel
2000 Nasal Reduplication in Mbe Affixation. *Phonology* 17: 65–115.

Wetzels, Leo and Joan Mascaró
2001 The typology of voicing and devoicing. Language 77: 207–244.

Regional variation in intonation: Conversational instances of the “hat pattern” in Cologne German[1]

Pia Bergmann

1. Introduction

The present paper deals with tonal variation on the suprasegmental level. The main objective of this study is to introduce the form and functions of a rising-falling intonation contour in the regional variety of Cologne German, spoken in the West Central German dialect area. Within the last few years, the study of regional variation of intonation has largely increased, mainly focussing on formal aspects of description, e.g. intonational inventories of different varieties or phonetic detail in the realization of different tonal categories such as accent types (cf. among others Grabe et al. 2000, Grabe and Post 2002, Gilles and Peters 2004, Barker 2005, Fletcher, Grabe, and Warren 2005, Gilles 2005, Kügler 2007, Peters 2006). Functional aspects are dealt with to a smaller extent, but are incorporated into the accounts of Peters (2006), Gilles (2005), and Kügler (2007), as well as Bergmann (2008).

Following Gilles (2005) the present study describes dimensions of regional variation at phonetic, tonological, and phonological levels, thereby including formal as well as functional aspects. The phonetic level of description deals with gradient details of phonetic realization, largely irrespective of the theoretical background of intonation analysis. The tonological level, on the other hand, includes a description based on autosegmental metrical theory (e.g. Ladd 1996). The intonation contour is discussed in terms of tonal targets, which are connected with each other by interpolation.[2] The main objective of the tonological description therefore is the allocation of established tonal categories to these tonal targets, which are assumed to evoke the excursion of fundamental frequency. Finally, the phonological level of description deals with the functions of a specific intonation contour, i.e. the way it is used by participants in an ongoing interaction.[3] The main concern of this three-fold description is to make the account a good basis for comparisons with similar intonation contours in other varieties, which might differ with respect to phonetic detail, tonological interpretation and/or phonological interpretation (Gilles 2005).

Along these lines the present study will first introduce the rise-fall and give an overview of its phonetic variation (section 3). Second, the tonological status of tonal targets will be discussed (section 4), and finally, it will be shown how the participants of an interaction make use of the rise-fall in the course of their conversation (section 5).

Since the rise-fall shows some formal similarity to the “ hat pattern” (see section 3), the phonological interpretation of the rise-fall will be carried out against the background of typical occurrences assumed for the “ hat pattern” in Standard German. It will be shown that the contours indeed have important characteristics in common concerning their usage, but that there are some peculiarities in the Cologne German rise-fall which contradict the “prototypical” usage of the “ hat pattern”. First, many instances of the Cologne German rise-fall do not conform to the partition of the rise-fall bearing utterances into a topic and a focus constituent (cf. Féry 1993, Mehlhorn 2001). Second, many of the rise-fall bearing utterances occur with a syntactic structure that is not mentioned for occurrences of the Standard German “ hat pattern”: It will be argued that an important shared feature of the Cologne German rise-falls lies in their impact on turn-taking and their potential for integrating two information units into one prosodic unit, i.e. intonation phrase. In light of these findings, the consequences of the underlying type of data (spontaneous dialogue vs. experimental/introspective data) will be discussed as well as the possibility of having a pitch accent on non-rhematic constituents. In the next section, the relevant facts about the data and methods of the study will be presented.

2. Data and methods

The present analysis is based on about 14 hours of spontaneous dialogue of dialect speakers from Cologne. The data comprise 6 interviews with elderly male and female speakers, 8 episodes of a half-documentary serial about a working-class family (“Die Fußbroichs”), and some episodes of the reality-TV show “Big Brother”. The TV serials were broadcasted in the 1980s and 1990s; the interviews were recorded in the year 2001. Out of this corpus, 51 instances of the rising-falling intonation contour have been chosen randomly. They have then been submitted to phonetic analysis with Praat (Boersma and Weenink 2006). The acoustic analysis was supplemented by auditive analysis. After that the formal and functional analysis was carried out.

The tonal analysis was conducted within the autosegmental metrical framework (e.g. Ladd 1996) where pitch accents are associated with metrically prominent syllables and boundary tones with edges of intonation phrases. In addition, we apply the analysis of the phrase tone (accent) that marks the edge of an intermediate phrase and is secondarily associated with a secondarily stressed syllable (cf. Beckman and Pierrehumbert 1986, Grice, Ladd, and Arvaniti 2000, Grice and Baumann 2002, Ladd 1996).

The functional analysis is couched into the framework of Interactional Linguistics (cf. Couper-Kuhlen and Selting 1996, Selting and Couper-Kuhlen 2000). intonation is viewed as a resource which can be employed by the participants to organize an ongoing interaction. In this respect, intonation may be considered a contextualisation cue, working as a signal to guide the interpretation of an ongoing interaction (cf. Auer 1992). Main methodological presumptions of Interactional Linguistics are (1) that the analysis should crucially rely on the observation of the participant's perspective, and (2) that the analysis should incorporate a detailed analysis of co-occurring linguistic devices, because resources on different linguistic levels like syntax, semantics, and prosody work together in guiding the interpretation of an ongoing interaction. The present study will give an account of the syntactic structures with which the pitch contour occurs, while at the same time taking into account actual responses of the participants of an interaction. Additionally, questions of information structure will be considered.

The combination of a formal description within the framework of autosegmental metrical theory and a functional interpretation within the framework of Interactional Linguistics is well suited to cope with the demands of a detailed analysis of regional intonation in naturally occurring dialogue. While the formal description is best done within a widely acknowledged framework in order to lead to results on which comparisons across different varieties may be based, the functional description must be able to come to terms with the often very complex and variable structures occurring in spontaneous dialogue. A drawback of the in-depth analysis of a rather small corpus of interactional data may be seen in the fact that generalizations about form and functions of intonation are difficult to achieve, since many influencing factors governing variation cannot be controlled. The advantages of the procedure however may make this a worthwhile undertaking: Having a close look at naturally-occurring dialogue within the framework of Interactional Linguistics may lead us to a better understanding of the whole range of variation interactants have to cope with in "real" conversation. The achieved findings can then be used to guide further research on the basis of a larger corpus.

3. Phonetics of the rise-fall

This section gives an account of the phonetic realization of the rising-falling contour. Before going into detail with the phonetic variation, the contour shall shortly be introduced, including some uncontroversial tonological interpretations: The contour consists of two pitch accents, the first one of which is a rising pitch accent and the second a falling one. In this respect, the pattern resembles the so-called "(flat) hat pattern" which is found in Standard German and other languages (cf. Féry 1993, Ladd 1996). The following picture illustrates the rise-fall in Cologne German. It occurs with the utterance "ob dat NARrenfreiheit is WEIS=isch nisch" (*Whether this is what one might call jester's licence, I don't know.*).

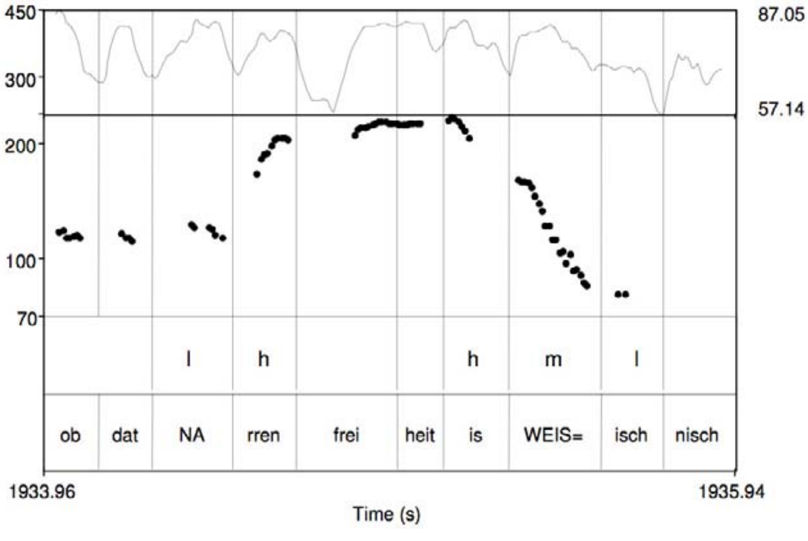


Figure 1: Rising-falling pitch contour in the regional variety of Cologne German.

The first pitch accent of the utterance is associated with the syllable "NA", the second pitch accent with the syllable "WEIS". The symbols on the tone tier give an approximate description of the pitch contour on a phonetic level with "l" referring to a low level of fundamental frequency, "h" to a high level and "m" to a mid level. They are not supposed to represent a tonological interpretation (cf. endnote 2 for notation conventions). It can be seen that the rising and falling pitch accents are connected by a high plateau thereby forming a prototypical instance of the pattern referred to as a “flat hat”.

A stylized representation of the rise-fall is given in the following picture. The boxes stand for syllables, the line symbolizes the pitch contour. Grey shaded boxes represent metrically strong syllables.

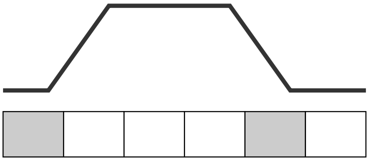


Figure 2: Cologne German rise-fall variant 1: Steep rise to high plateau.

Another variant of the rise-fall differs from this one by a gradiently falling fundamental frequency instead of the high plateau, as can be seen in the following picture (fig. 3).

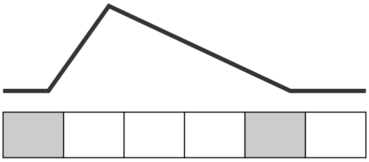


Figure 3: Cologne German rise-fall variant 2: Steep rise to high peak.

The alignment of the peak can vary, so that realizations with a high peak that is not reached until two syllables after the accented syllable exist as well (fig. 4). These realizations are termed variant 3.

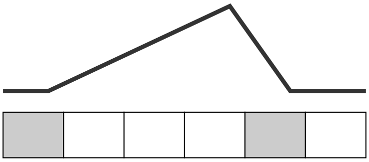


Figure 4: Cologne German rise-fall variant 3: Slow rise to high peak.

In addition to these variants another realization of the rise-fall can be found. It is characterized by a slow rise to a high plateau (fig. 5).

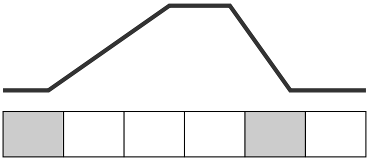


Figure 5: Cologne German rise-fall variant 4: Slow rise to high plateau.

Variants 1-3 are captured in the account of Cologne German intonation of Peters (2006) too. Variant 4 on the other hand is not incorporated into Peters' account. In our corpus there is only one instance of a high plateau after a slow rise. The rare occurrence of this variant in the investigated corpus and its absence in Peters (2006) may be due to the fact that the relevant segmental stretch of the rise-fall bearing utterances in many cases simply is not long enough to allow for a high plateau after a slow rise. Other variation than that of the rising movement and the alignment of the high peak or plateau concerns the level of the first pitch accent, which may be low or mid.

In addition to the variation of the first, rising part of the rise-fall different realizations of the falling part can be observed as well: The second pitch accent shows considerable variation concerning its height. It may be realized on mid, low or high pitch level. In most of all cases (28 out of 50)[4] the second pitch accent isrealized on mid pitch level. It should be noted, however, that this preference for the realization on mid pitch level may be a simple effect of the position of the second pitch accent within the falling part of the rising-falling contour, i.e. the earlier the accented syllable occurs within the second part of the rise-fall bearing utterance, the higher its fundamental frequency will be and vice versa. Thus, the following rise-fall bearing utterances exemplify a second pitch accent that occurs late after the pitch peak (fig. 6) and one second pitch accent that occurs early after the pitch peak (fig. 7), respectively. According to their position within the overall falling movement the late pitch accent has low pitch level, whereas the early pitch accent is realized on high pitch level.[5]

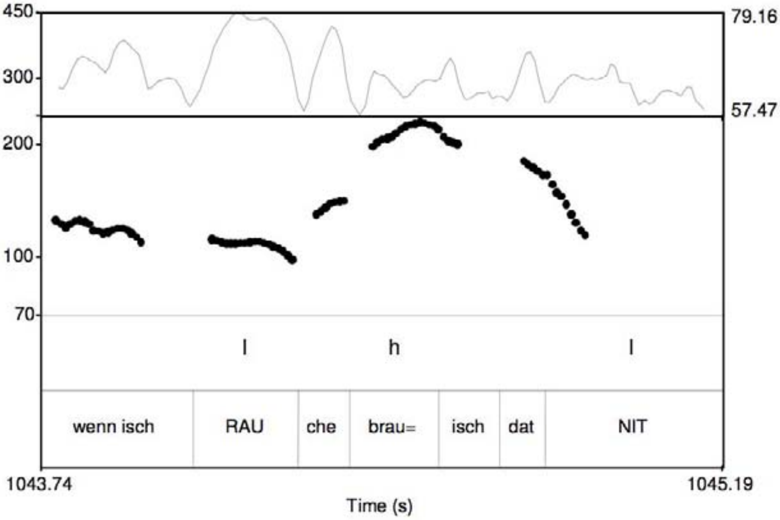


Figure 6: Second pitch accent late after the high peak ('When I smoke I don't need that').

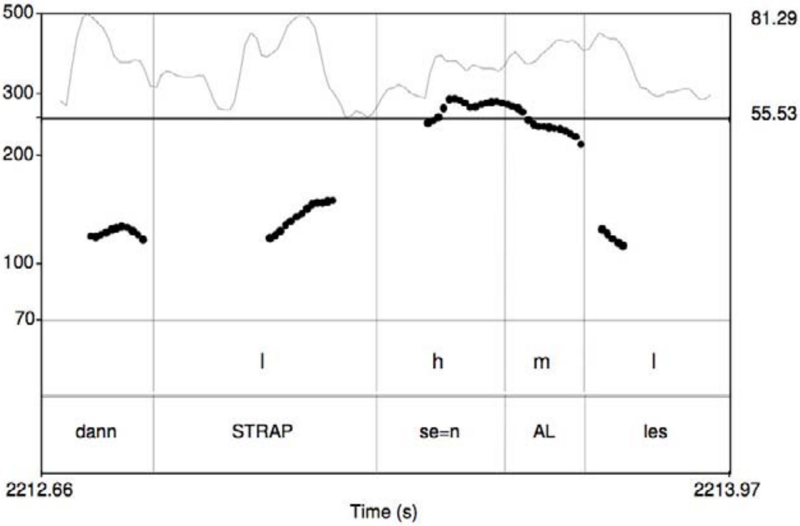


Figure 7: Second pitch accent early after the high peak ('Then suspenders and everything').

For the pitch accent to be produced on a low level a distance of at least three syllables is necessary between the pitch peak and the second pitch accent (fig. 7).

Another aspect of the varying falling pitch movement concerns the realization of the final low. Given enough space and time after the second pitch accent, most of the rising-falling contours are characterized by a late low plateau. For tonological discussion of this fact see section 4.2.

The following illustration serves as a summary of the variation in the Cologne German rise-fall.

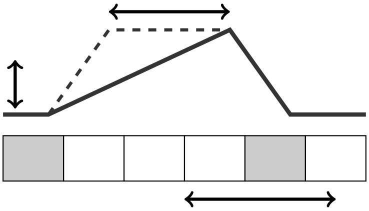


Figure 8: The Cologne German rise-fall and its phonetic realization.

From left to right it can be seen that firstly, the first pitch accent varies with respect to its height, ranging from low to middle. Secondly, the rise to high may be quick or slow, i.e. the alignment of the high peak varies. Thirdly, the rise to high may result in a pointed peak or a plateau, which is indicated by the dotted line. Finally, the height of the second pitch accent varies depending on the location of the metrically strong syllable with respect to the falling movement of the rise-fall.

4. Tonology of the rise-fall

The next section deals with the tune-text-association of the rising-falling contour. First, the tonology of the rising movement will be considered (section 4.1.). Second, we will discuss the tonology of the falling movement (section 4.2.).

4.1. Tonology of the rising movement

The section on phonetic variants of the rise-fall has firstly demonstrated that the beginning of the rise starts in the accented syllable and secondly that the alignment of the high peak or plateau may vary. Since the rise consistently starts in the accented syllable we assume a tonological low target (l*) associated with the accented syllable.

As for the varying pitch peak it may be located on the first syllable after the accented syllable (fig. 2, 3), but it may also be located considerably later (fig. 4, 5). In the following, I will discuss two possible explanations for the alignment of the high peak. First, it will be considered whether the high pitch peak may be interpreted as a trailing tone of the low pitch accent l*. Second, it will be discussed, if the high peak can be interpreted as a high boundary tone. Both explanations have in common that they allow for the alignment of the high peak with an unstressed syllable, which is the case for many of the high peaks in the rise-fall under discussion.

In order to clarify the tonological status of the high peak, we will first consider the following two examples. Position of small letters indicates the alignment of tonal targets.

(1)

			l h	l
sie	*is*	*die*	*BESte*	*momenTAN*
she	is	the	best	at the moment

‘She is the best at the moment.’

(2)

				l		h	h	m l	
aber	*wir*	*müssen*	*natürlisch*	*WARten*	*wat*	*mit*	*de*	*OMma*	*is*
but	we	must	of course	wait	what	with	the	Grandmother	is

‘But we have to wait of course and see what happens with Grandmother.’

In the first example the h-tone is aligned on the post-accentual syllable, whereas the second example is characterized by a later peak-alignment. The alignment in (1) argues in favour of a high trailing-tone, which by definition should be aligned near the pitch accent it is part of (cf. Grice and Baumann 2002, Ladd 1996). The late alignment of the peak in (2) on the other hand causes a problem for this interpretation of the high tonal target. Despite this variable alignment of the high peak, Peters (2006) considers the tonal target of the peak a trailing tone. This is due to the fact that in his account trailing tones are unassociated tones with variable alignment. I.e. for a trailing tone, it is not necessary to align in close distance to a pitch accent. In the present study this interpretation will finally be adopted. The remainder of this section will show why this interpretation is superior to the interpretation as a high boundary tone.

First of all, the auditive impression of the rise-fall does not suggest the division of the contour into two parts.[6] Nevertheless, a point of interest with respect to the interpretation of the h-tone as a boundary tone lies in the findings of Truckenbrodt (2005). He analyses the occurrence of upstepped pitch as a signal for edge marking in embedded sentences, and asserts that "the right edge of embedded clauses, though not the left edge, leads to an intonation phrase boundary" (Truckenbrodt 2005: 273). Since many of the Cologne German rise-falls have an extraordinary high peak, there is one similarity between Truckenbrodt's data and our pitch contour. Another parallel between the Cologne German rise-fall and the sentences investigated by Truckenbrodt concerns the occurrence of a syntactic boundary: More than 80% of all rise-fall bearing utterances are characterized by a syntactic boundary. Thus the question arises whether syntactic phrasing influences intonational phrasing, as Truckenbrodt assumes for his examples of upstepped h-tones. To judge this, all instances of the rise-fall with a syntactic boundary (n=41) were analysed with respect to their tune-text-association of the high peak. If the high peak should be invariably aligned with the syntactic boundary, this surely might hint at the possibility of interpreting the h-tone as a boundary tone. The results of the analysis are not clear-cut, however. In contrast to the findings of Truckenbrodt no consistent alignment of the h-tone with the edge of a syntactic constituent could be detected. Taking the right (or final) edge of the first syntactic unit as the target for the alignment of the h-tone, the analysis yields the result that this is the case for only 14 instances (34.1%) of all rise-falls with a syntactic boundary (see e.g. (1)). 19 instances (46.3%) of all cases have the h-tone aligned with the left or initial edge of the following syntactic unit, thereby producing a more or less continuous pitch movement across the syntactic boundary, as in (3). (Syntactic phrasing is marked by square brackets):

(3)

		l	h	l
[du	*willst*	*TANzen]*	*[heut*	*ABend]*
you	want-2PsSg	dance	today	evening

'You want to go dancing tonight.'

Another 5 rise-falls (12.2%) show a high plateau across the syntactic boundary, i.e. the right edge as well as the left edge of the subsequent syntactic unit is realized on a high pitch level:

(4)

	l h	h		l
[vierzehn	*TAge]*	*[is*	*dat*	*SCHÖN]*
fourteen	days	is	that	nice

‘That is nice for a fortnight.’

The beginning of the high plateau starts on the last syllable of the syntactic constituent „vierzehn TAge“ (‘for a fortnight’) in the so-called front-field (*Vorfeld*) of the sentence; the fall starts with the beginning of the inner-field (*Mittelfeld*) on the syllable “dat” (‘that’). (The term front-field (*Vorfeld*) refers to the position in front of the finite verb in German sentences, inner-field (*Mittelfeld*) refers to the position after the finite verb up to the infinite verb, if there is one. The position after the infinite verb is called end-field (*Nachfeld*), cf. Auer 1996). Finally, in 3 instances (7.3%) of all rise-fall bearing utterances with a syntactic boundary the h-tone is not reached until the middle of the second syntactic unit:

(5)

					l		h l
[un	*wer*	*hat*	*sisch*	*die*	*überLEGT]*	*[die*	*PREIse]*
and	who	has	REFL	PRO-ACC	think	the	prices

‘And who thought of these, the prices?’

The h-tone is aligned with the accented syllable of the right dislocated constituent „die PREIse“ (‘the prices’).

To summarize, the h-tone has a tendency towards the alignment with a syntactic edge in 33 out of 41 rise-falls. The alignment is rather inconsistent, however, and alternates between the right edge of the first syntactic unit and the left edge of the second syntactic unit. Of the remaining 8 rise-falls, 5 have a plateau across the syntactic boundary, and 3 do not align the h-tone any earlier than in the middle of the second syntactic unit.

Apart from the above discussed rise-falls, our corpus contains 10 rise-falls without any syntactic boundary. The following example is one of these instances:

(6)

			l	h	m l
[der	*war*	*doch*	*GRAde*	*hier*	*in der Nähe]*
PRO-NOM	be-past	PART	just	here	close by

'He was just here close by.'

The rise-fall on this utterance does not differ from those on utterances with a syntactic boundary. The h-tone is reached two syllables after the low pitch accent, the fundamental frequency then falls until the second pitch accent on mid level is reached. Taking into account these rise-falls without a syntactic boundary still reduces the proportion of rise-fall bearing utterances, so that the h-tone is aligned at a syntactic boundary in 64.7% all in all.

To sum up, it may be stated that the variable alignment of the h-tone with respect to a syntactic boundary as well as the lack of other cues to an intonation phrase boundary (c.f. note 6) argue against the interpretation of the h-tone as a boundary tone. Consequently, the tonal status of the h-tone will be considered an unassociated high trailing-tone, which is in line with the interpretation of Peters (2006). The differences in alignment are allocated to phonetic implementation. The high proportion of occurrences with syntactically complex utterances, however, is still a peculiarity of the rise-fall in Cologne German, which will be discussed in section 5.

4.2. Tonology of the falling movement

In order to elucidate the tonology of the falling movement it is necessary to clarify the status of the second pitch accent as well as the status of the pitch movement following the second accent and leading to a final low until the end of the intonation phrase. Section 3 has introduced different phonetic realizations of the second pitch accent. It has been mentioned that the scaling of the pitch accent tends to be correlated with its position within the falling movement. The final low usually is reached with some distance to the intonation phrase boundary, thereby producing a low plateau.

Many realizations of the second pitch accent suggest the tonological interpretation of a downstepped h-tone. Downstep is commonly seen to be a phenomenon which only occurs after another h-tone. In addition to pitch accents, other tonal categories like trailing-tones, phrase accents and boundary tones may be submitted to downstep (cf. Grice and Baumann 2002). The pitch accent in the rising-falling contour follows a high trailing-tone and could therefore possibly be subject to downstep. The phonetic realization of the downwards movement is indeed very similar to the one described by Grice and Baumann (2002) for downstepped accents in Standard German (GToBI). The GToBI system includes a downstepped accent which is subsequent to a high leading-tone (H+!H*). It is characterized by a pitch accent on mid intonation level, itself not evoking a pitch excursion. Thus, comparably to the Cologne German rise-fall, the pitch accent is part of a continuously falling movement. Contrary to Standard German, however, the beginning of the falling movement in Cologne German is due to a sequence of a high trailing-tone and a downstepped high pitch accent: +h !h*. This tonological interpretation holds for a large part of the rising-falling contours. Those instances of the rise-fall with a non-downstepped second pitch accent as well as those instances with a low second pitch accent pose a problem, however.

First, concerning the rise-falls with a non-downstepped pitch accent, 6 out of 51 rise-fall bearing utterances fall into this group. The high peak of these instances is not reached until the second pitch accent, which is not downstepped with respect to the trailing tone of the first pitch accent. An example is the following contour bearing utterance:

(7)

	l h	h l
[wills=e	*VORjehn]*	*[ANnemie*
want-2PsSg-you	go ahead	NAME

‘Do you want to go ahead, Annemie?’

In these cases the h-tone has to be considered an “ordinary“, non-downstepped h-tone: h*. All in all, those rise-falls with a downstepped second pitch accent clearly prevail compared to those without downstep.

Second, for the rise-falls with a pitch accent on a low level the question arises whether downstep may be as strong as to result in low pitch level. Concerning this question, the findings of Grabe (1998) are of some interest: She found that “partial downstep” as well as “total downstep” exist in a variety of Northern Standard German (Braunschweig). The notion “total downstep” refers to a downstepped pitch accent, which is realized on the level of the final low boundary tone. Principally, the pitch level of the downstepped pitch-accent seems to be characterized by continuous variation, i.e. no discrete bundles centre around either low or mid level realizations, but there are various in-between realizations. Consequently, low and mid level realizations are not assumed to be categorically distinct, but are subsumed under the notion of downstep. Although this finding should be tested on a larger data base for the Cologne German rise-falls, it shall be assumed that here, too, the category of downstep may be transferred to the pitch accents on the low intonation level. An example of a completely downstepped pitch accent in a rise-fall is given in (8):

(8)

		l	h	l
[du	*willst*	*TANzen]*	*[heut*	*ABend]*
you	want-2PsSg	dance	today	evening

‘You want to go dancing tonight?’

To sum up, in most cases the second pitch accent may be interpreted as a partially downstepped h-tone, which is realized in line with a gradiently falling downwards movement (!h*). Few instances have been considered to bear total downstep. Another marginal tonological variant of the rise-fall is characterized by a non-downstepped pitch accent (h*).

The remainder of this section deals with the tonology of the final low, which is reached early in the intonation phrase, so that a final low plateau can be observed in many cases. This alignment suggests a low phrase accent as a tonal category of the l-tone. Unfortunately, the alignment of the l-tone can be tested for in only 20 rise-falls, since the remaining instances of the contour simply are not long enough to leave a choice for the alignment of the tone. With the exception of 5 instances, all out of 20 rise-falls align the l-tone earlier than on the last syllable of the intonation phrase. Consider for example the following two rise-fall bearing utterances (9-10). These align the beginning of the final low with the secondarily stressed syllable after the second pitch accent. As has been mentioned in section 2, according to Grice, Ladd, and Arvaniti (2000), this is the expected anchor for the alignment of a phrase accent (only the stretch with the falling movement is given in the following examples):

(9)

		m	l
[musst=e	*de*	*HEIzung*	*ausdrehn]*
must-2PsSg-you	the	heating	turn off

‘you have to turn off the heating.’

(10)

				m	l
[bis	*es*	*an*	*der*	*TÜR*	*ruskam]*
until	it	at	the	door	came out

‘until it came out of the door.’

In five further instances the secondarily stressed syllable coincides with the last syllable in the intonation phrase, so that it cannot be determined whether the low tone should be interpreted as a phrase accent or a boundary tone of the IP, an example of which is given in (11):

(11)

	m	l
[die	*KNEIpe*	*da]*
the	pub	there

‘the pub there.’

If there is no more secondarily stressed syllable in the intonation phrase, the low pitch is aligned with the end of the IP:

(12)

	m l
[aus	*THÜringen]*
from	Thuringia

‘from Thuringia.’

Therefore, the final low of the Cologne German rise-fall will be described as the tone sequence l-%. The complete falling movement is thus assumed to consist of two possible tone sequences: +h !h* l-%, and +h h* l-%.

Concluding this section, the tonology of the Cologne German rise-fall may be characterized by the following two tone sequences:

(13)

a. l*+h !h* l-%
b. l*+h h* l-%

(13a) stands for those instances of the rise-fall with a downstepped second pitch accent, while (13b) characterizes those instances without downstep. (13b) occurs only rarely compared to (13a).

5. Occurrences of the rise-fall in natural conversation: a typical “hat pattern”?

This section aims to describe the typical conversational instances of the rise-fall in Cologne German and thereby identify the function of the rise-fall. One central aspect of the account will be the syntactic structure of the rise-fall bearing utterances which has been touched upon in section 4.1. In addition to this, the function of the rise-fall is discussed with respect to information structure. As has been mentioned in section 3, the Cologne German rise-fall resembles a contour commonly referred to as “ hat pattern”. It will be discussed whether the formal resemblance is mirrored by a functional similarity, and it will be demonstrated that some of the conversational instances of the Cologne German rise-fall show considerable overlap with the assumed functions of the “ hat pattern” in Standard German, but there are several instances which do not fit into the picture. Both will be demonstrated below. First, we will introduce those rise-falls that do not deviate from the Standard German rise-fall (section 5.1). After that the deviations from the Standard German rise-fall will be illustrated (section 5.2).

5.1. Functional conformity with the “hat pattern” in Standard German

Functional interpretations of the Standard German rise-fall or “ hat pattern” assume its partition into a topic constituent which occurs with the rising pitch accent of the contour and a focus constituent which occurs with the falling pitch accent (cf. Féry 1993). Mehlhorn (2001) states that the rising accent on the topic constituent serves to make the „neutral“ topic more prominent, a procedure also referred to as i-topicalisation (cf. Jacobs 1997, Mehlhorn 2001). According to Jacobs and Mehlhorn, it is particularly important in this respect that the constituent in topic-position builds up a contrast with another element, which may be explicit or implicit. Several syntactic structures are supposed to be typical contour bearing structures, including utterances with qantifiers and discontinuous constituents (cf. Mehlhorn 2001), as well as utterances with multiple foci, gapping-constructions, and certain types of complex sentences (cf. Féry 1993).

Many of the Cologne German rise-falls from the investigated corpus do not deviate from the Standard German rise-falls. An example is given in (14). It demonstrates a conversational instance with a rising pitch accent on the topic constituent and a falling accent on the focus constituent, just as it might be expected considering the description of the “ hat pattern” in Standard German. The context of the utterance is a conversation about possibly negative effects of playing tennis.

(14) k02-schläjer ja

```
     13   fm   ah doch die sin doch LEISCHT; oder,
               ‘But they are light, aren’t they’
     14   fv   WAS.
               ‘What’
     15   fm   <<p>die SCHLÄ[ger?>
               ‘The rackets’
                       l h  hl
⇒    16   fv   [die SCHLÄjer JA.
               the  rackets  yes
               ‘Yes, the rackets’
     17        aber die WUCHT wo der BALL kommt.
               ‘but the force of the ball’
     18        wo de den zuRÜCK schlägst;
               ‘at the point where you hit back’
     19        das jeht ja alles in die jeLENke un:-
               ‘this damages the joints an-’
```

The example is an instance of a gapping-construction (line 16), which refers back to fm’s utterances in lines 13 and 15. The rising movement stretches across the constituent “die SCHLÄjer” (‘the rackets’), while the falling movement covers the word “JA” (‘yes’), which is the focus of the utterance. The short sequence following the utterance under discussion exemplifies the potential contrast incorporated in the rising-falling utterance. In other words, the contrasting element to “die SCHLÄjer” is made explicit by fv, who introduces the element “die WUCHT” (‘the force’) by the adversative conjunction “aber” (‘but’). Thus, as assumed by Mehlhorn on the basis of isolated, invented sentences, the topic-constituent with a rising movement evokes a contrasting element, which in this case is indeed made explicit in the next utterance. This is true for four other instances of the rise-fall. In all cases, the contrasting element is made relevant in the ongoing interaction through the participants. Further instances, although not necessarily conforming to the partition into a topic and a focus constituent, support other functions of the Standard German rise-fall, e.g. the occurrence with multiple foci or complex sentences (for a complete account of all utterances see Bergmann 2008).

A shared feature of all instances in this group is the fact that they include two accented information units. These information units are integrated into one prosodic unit by the rising-falling contour. Peters (2006) asserts the same effect for a Northern Standard German intonation contour, consisting of a prenuclear H*-accent and a nuclear H*L-accent. These contours, which are rising-falling on the surface, too, serve to present two information units as part of a whole. While Standard German has a high pitch accent on the first constituent, it is low (or rising) in Cologne German. Please note that the level of the second pitch accent shows internal variation in the Cologne data. Most of all examples have a second pitch accent on mid level (interpreted as !h*), but there exist instances with a high (h*) pitch accent, or with a pitch accent on low level, which was assumed to be subject to total downstep (see section 4.2).

In this section it was demonstrated that some of the Cologne German "hat patterns" resemble the Standard German hat patterns with respect to their information structure and to their syntactic structure.

5.2. Functional deviations from the "hat pattern" in Standard German

It has been pointed out several times that the rise-fall in Cologne German also occurs with utterances in which the " hat pattern" does not typically occur. 21 instances out of 51 connect the falling accent with a constituent, which expands a syntactically complete sentence. The notion of "expansion" of a syntactically complete utterance originates from conversation analytic research.[7] It is based on the observation that in naturally occurring conversation, syntactic units which are complete in the sense of not projecting any further constituents to occur are often extended by other constituents, as e.g. in (1) and (3) above.

The expanding constituents are dependent on the previous structure either syntactically, semantically, pragmatically and/or prosodically.[8] Since in the present study all examples under discussion have a similar prosodic structure – no intonation phrase boundary separating the rising and the falling part of the contour, but having one pitch accent in each part – the way the extending constituent depends on the previous structure prosodically does not vary. There is some variation concerning the syntactic dependence, on the other hand, which will be taken as a starting point of the following description. Furthermore, the account aims to shed some light on the information structure of the expanding constituent. The main objective of this section is to prove whether or not the rise-fall bearing utterances under discussion deviate from the assumption that the falling accent is connected with a focus accent. Following Uhmann (1991), the focus of an utterance will be considered the part which is prominent or newsworthy with respect to content. The terms "focus" and "rhematic information" will be used interchangeably. With the terms "topic" or "non-rhematic information" I refer to those parts of the utterance, which are not prominent with respect to content, because they are given or inferable.

According to Auer (1991, 1996) syntactic expansions may be either regressive or progressiv. While a syntactic structure is modified by regressive expansions, it is continued by progressive expansions. Regressive expansions are of two types, which differ with respect to the insertability of the expanding constituent into the previous syntactic structure. Whereas syntagmatic regressive expansions may retrospectively be inserted into the previous structure, this is not the case for paradigmatic regressive expansions. These substitute an element of the previous structure. In the investigated corpus, 7 rising-falling utterances belong to the progressive type, 9 instances are syntagmatic regressive expansions, and 5 instances are considered paradigmatic regressive expansions. The following table summarizes the relevant facts. Expanding constituents are marked by underline, the position, where the expanding constituent could be inserted into the previous structure is marked by underlined X.

After this brief introduction, let us now turn to the information structure of the different syntactic types of rise-fall bearing utterances. I will start with the paradigmatic regressive expansion, then turn to the syntagmatic regressive expansion and will finally discuss the progressive expansions.

5.2.1. *Paradigmatic regressive expansions*

Since this type of expansion per definition replaces a previously introduced element, it can be assumed that the expanding constituent has no rhematic status. This assumption is confirmed by the 5 utterances in the corpus. According to Auer (2006) this type of expansion tends to be prosodically integrated into the previous structure, which of course is true for all instances under discussion. Example (15) shall serve as an illustration of this fact:

(15)

					l	h l
un	*wer*	*hat*	*sisch*	*die*	*überLEGT*	*die PREIse*
and	who	has	REFL	PRO-ACC	think	the prices

‘And who thought of these, the prices?’

The constituent “die PREIse” (the prices) expands the previous syntactic structure and substitutes the proform “die” of that structure. It thereby clarifies a previously introduced constituent; no new information is given nor is an assertion made. Thus, concerning the paradigmatic regressive type, the falling pitch accent cannot plausibly be considered a focus accent.

5.2.2. *Syntagmatic regressive expansions*

For the syntagmatic regressive expansions no consistency with respect to information structure can be claimed, as it has been done for the paradigmatic regressive expansions in the previous section. The majority of examples (n=5)[9], however, is considered non-rhematic, as demonstrated by (16). The utterance is subsequent to a reproach by the same speaker concerning the way her partner spends his spare time on weekends (*du kAnnst nisch immer WEGgehn WEGgehn*; ‘you cannot go out all the time’):

(16)

			l	h	m	h l
und	*dann X*	*nisch*	*traiNIERN*	*gehen*	*am*	*WOchenende*
and	then	not	work-out	go	on	weekend

‘And then not to go to the training on the weekend.’

The expanding constituent is inferable from the context and presumably not informationally prominent, since the previous utterances refer to weekend acitivities, too. Thus, no contrast between the constituent “am WOchenende” (‘on the weekend′) and another element is made. Therefore, the information is considered non-rhematic information, which is in line with the interpretation in (37) above.

In contrast, the remaining two instances may be considered rhematic as shown in (17):

(17)

				l	h	m l
die	*amriKAner*	*gingen*	*X*	*RAUS*	*aus: (-)*	*THÜringen*
the	Americans	go-3PsPlPast		out	from	Thuringia

‘The Americans withdrew from Thuringia.’

The context of this example is a longer contribution of the current speaker concerning her evacuation from and return to Cologne after World War II. The information unit “aus THÜringen” has not been mentioned before and is crucial for the understanding of the narrated events. For this reason the expansion is considered to bear rhematic information.

5.2.3. *Progressive expansions*

Finally, the progressive expansions will be demonstrated. Most of them contain semantically void information, as can be seen in the following example (18):

(18)

				l h		m l
da	*gAb=et*	*noch*	*keine*	*BLÄCKföß*	*oder*	*SOwat*
there	give-3PsSg-Past	yet	no	NAME	or	anything

‘The Bläckföß[name of a band, P.B.] or anything like that did not exist yet’)

The expanding constituent serves as a hedge marker and does not supply any rhematic information. In almost all instances of this group, the expanding constituent bears a pitch accent on mid level.

On the basis of the data presented we conclude that the expanding constituents do not introduce a rhematic element into the discourse. Therefore, the assumption that the falling pitch accent is connected with a focus constituent does not hold for these instances. The conversational usage of the Cologne German rise-fall therefore shows considerable variation concerning not only the syntactic structure of the carrier utterances, but also concerning their information structure. It may be concluded, then, that the rise-fall primarily serves as an integrating device, presenting two syntactically or informationally distinct entities as parts of a whole. The informational weight of the second “falling” part may vary. The syntactic completion point, which might indicate a possibly complete turn constructional unit (TCU), is not used for turn-taking, since prosodic devices mark the ongoing unit as being incomplete (cf. Selting 2000). Detailed analysis of the contextual embedding of the rise-fall bearing utterances may lead to further insights into the conversational functions of the rise-fall in Cologne German.

6. Summary and conclusions

The main aim of this study was to present a descriptive analysis of the rising-falling intonation contour in the regional variety of Cologne German. To this end, formal as well as functional aspects of the rise-fall were discussed. The formal description included a discussion of phonetic and tonological variation and interpretation, while the functional description focussed on the way the rise-fall is used in naturally occurring data. Here, the co-occurrence of the intonation contour with different syntactic structures was presented. The assumed occurrences of the so-called " hat pattern" in Standard German served as a background for the discussion of specificities of the Cologne German rise-fall.

All in all, we have shown that the rise-fall is subject to a rather large amount of internal variation on all levels of description. This is specifically true for the phonetic level with much variation concerning mainly the horizontal realization of the pitch peak or plateau. Tonological analysis yielded two variants, one of which is marked by downstep of the second pitch accent, whereas the other lacks this feature. Phonological (or functional) interpretation showed considerable overlap of the conversational usage of the contour as compared to the " hat pattern" in Standard German. This overlap includes those instances of the contour with multi-focal structure, or with a clear topic-focus-division, with the falling pitch accent being aligned with a constituent belonging to the focus of the utterance. Thus, the assumed usage of the " hat pattern" in Standard German, which is mostly drawn from introspection or experimental data, can be confirmed on the grounds of naturally occurring data for the regional variety of Cologne German.

In contrast to these occurrences of the rise-fall, however, many of the Cologne German rise-falls were characterized by a diverging occurrence pattern. First, considering syntax, it was demonstrated that the falling movement of the rise-fall often occurs with expansions across a syntactic completion point. This syntactic structure as a carrier sentence for the rise-fall is not mentioned in the relevant literature concerning the rise-fall. Second, with respect to information structure, it could be shown that the second pitch accent, which is !h* or h*, is related to constituents with widely diverging informational status. In other words, there are instances which associate the pitch accent with highly rhematic information, but there are other instances as well which associate the pitch accent with information of limited informational value. This finding is not in agreement with common descriptions of the Standard German rise-fall. While the first deviation from the Standard German " hat pattern" might be explained rather easily, the second deviation is more difficult to interpret. Thus, the absence of "hat patterns" with expansion to the right in the existing accounts might simply be due to the fact that this structure is typical for spontaneous conversation because it is intricately intertwined with upcoming demands of turn-taking and information processing in real time. The structure therefore is not likely to be produced for experimental settings or by introspection. In order to be able to decide whether this deviation is at least partly due to regional specificities of the Cologne German variety or solely due to the type of investigated data, comparisons with rise-falls in spontaneous dialogue of other varieties need to be undertaken.

The second deviation, on the other hand, raises more complicated questions. It may be questioned whether constituents without rhematic status may be associated with a pitch accent at all. Considering the assumed linguistic functions of pitch accents, namely prominence-lending, this clearly should not be the case (but see Baumann and Grice (2006) for a short introduction into this topic). Recent research on different degrees of "givenness" has shown that a binary distinction between "given" and "new" entities is not viable, and intonational accent-correlations for a three-way distinction have been found (cf. Baumann 2005). Constituents on the "givenness"-pole of the continuum, however, are still correlated with the option of taking "no accent". It has to be borne in mind that the notions of "givenness" or "newness" as Baumann uses them are not equivalent to the notion of "rhematic" information or focus drawn upon in this study. As Baumann points out, studies on the phenomenon of givenness often differ with respect to the definition of givenness. He identifies three main notions the term is applied to: identifiability, degree of activation, and focus-background structure. While the first and the second notion refer to cognitive states the respective entity has in the mind of the „listener“, the third one stands for the pragmatic role the entity has, i.e. it may be marked as "newsworthy" or "not newsworthy" by the speaker. Whereas Baumann's study is centred on the second notion, the present study focuses rather on the notion of "newsworthiness" or informational „importance“ of certain constituents. Still, the intonational option of “no accent” for non-rhematic status would probably be more plausible than the marking by a !h*-accent, which is found in the present data. This is especially true in the face of the iconic use of intonation, which Baumann claims to be valid for the prosodic marking of different degrees of givenness (concerning the iconic use of intonation see e.g. Ohala 1983, Gussenhoven 2004). The empirical evidence presented in our study, however, calls into question the invariant correlation of accentuation with rhematicity. More research on the basis of a larger corpus is necessary to resolve this issue.

Another aspect of the presented data which is not in agreement with the findings of Baumann concerns the correlation of accent type with accent degree. Not only is the Colognian rise-fall characterized by the peculiarity of associating a pitch ‘accent’ with non-rhematic information, but the same accent type (!h*) is also associated with different degrees of rhematicity. Of course, the correlation is not supposed to be an all-or-nothing manner, but the present study does not even suggest a tendency towards such a correlation. This is in line with Féry (2007), who presents further evidence against the invariant correlation of intonational elements with information structural categories. Finally, the diverse usage of the !h*-accent concerning the information status of the associated constituent calls into question major assumptions of the compositional approach of intonational meaning, where the accent type is supposed to have a “fixed” meaning (see e.g. Pierrehumbert and Hirschberg 1990). The data therefore speaks in favour of a holistic interpretation of the rise-fall, not necessarily aligning the second pitch accent with a focus constituent in all instances of the contour, but leaving space for another common feature of the rise-fall bearing utterances: its impact on turn taking, i.e. turn-holding across a syntactic completion point and/or the integration of two informational units into one prosodic unit.

Notes

1. For many helpful comments on my paper I thank an anonymous reviewer and Frank Kügler.
2. Tones on the tonological level are represented as small letters with discritics indicating their tonal status, e.g. l*, whereas the tones on the phonological level are represented as capital letters with diacritics, e.g. L*. Small letters without diacritics refer to the phonetic level.

3. Our analysis of intonational functions focuses on conversational data. Attempts of modelling intonational meaning within the autosegmental metrical framework are thus left aside since the object of study is different. This becomes clear in section 2, where the data of the present study is explained in more detail.
4. One instance of the rise-fall could not be analysed acoustically due to overlapping talk.
5. There is one counter-example, where the second pitch accent follows immediately after the high peak and obviously triggers a jump to low. In this case no obedience towards the tendency to realize the pitch accent in line with a gradient downwards movement can be observed.
6. In general, the division into intonation phrases still is a problematic issue in intonation research. Several cues are supposed to contribute to the phrasing of intonation contours such as pitch jumps on unstressed syllables, final lengthening, the occurrence of silence, anacrustic syllables and/or parallel intonation movements in a sequence (cf. Grabe 1998). Each intonation phrase should contain minimally one pitch accent (cf. Pierrehumbert 1980, Cruttenden 1997). With regard to these cues only pitch jumps on unstressed syllables occur in some cases of the rising-falling contour. There are no parallel structured intonation movements in a sequence, since the peak or plateau divides the movement into a rising and a falling part. Although final lengthening should be tested for by means of a detailed acoustic analysis, on auditive grounds no tendency for either final lengthening or anacrustic syllables could be detected. The same holds for silences between the rising and the falling part of the movement. Thus, evidence in favour of an intonation phrase boundary-tone seems to be scarce for the rise-fall under discussion.
7. For a critical survey of existing concepts of the notion of expansion (“increment” or “extension” respectively) see Auer (2006).
8. As Auer (2006) emphasises, the question of how the extending constituent depends on its previous structure is not yet decided. Especially, it is not clear how the dependencies on the different linguistic levels are structured with respect to each other. Therefore, Auer proposes to take the syntactic structure of the units as a starting point, and at the same time to give a detailed account of at least the co-occurring prosodic and pragmatic structure. Furthermore, the author hints at the fact that gestural details might be of interest as well.
9. Two of nine instances of this type cannot be futher analysed due to a cut in the data right before their beginning.

References

Auer, Peter
1991 Vom Ende deutscher Sätze. *Zeitschrift für Germanistische Linguistik* 19: 139–175.

Auer, Peter
1992 Introduction: John Gumperz' approach to contextualization. In *The contextualization of language*, Peter Auer, Aldo di Luzio (eds.), 1-38. Amsterdam: John Benjamins.

Auer, Peter
1996 On the prosody and syntax of turn continuations. In *Prosody in conversation,* Elizabeth Couper-Kuhlen, Margret Selting (eds.), 57-100. Cambridge: Cambridge University Press.

Auer, Peter
2006 Increments and more. Anmerkungen zur augenblicklichen Diskussion über die Erweiterbarkeit von Turnkonstruktionseinheiten. In *Grammatik und Interaktion,* Arnulf Deppermann (ed.), 279-294. Radolfzell: Verlag für Gesprächsforschung.

Barker, Geoffrey
2005 *Intonation patterns in Tyrolian German. An autosegmental-metrical analysis.* New York: Peter Lang.

Baumann, Stefan
2005 Degrees of givenness and their prosodic marking. *Online-proceedings of the International Symposium on "Discourse and Prosody as a complex interface",* Aix-en-Provence.
http://aune.lpl.univ-aix.fr/∼prodige/idp05/actes/baumann.pdf

Baumann, Stefan, and Martine Grice
2006 The Intonation of Accessibility. *Journal of Pragmatics* 38: 1636-1657.

Beckman, Mary E., and Janet Pierrehumbert
1986 Intonational structure in Japanese and English. *Phonology Yearbook* 3: 255-309.

Bergmann, Pia
2008 *Regionalspezifische Intonationsverläufe im Kölnischen.* Formale und funktionale Analysen steigend-fallender Konturen. Tübingen: Niemeyer.

Boersma, Paul, David Weenink
2006 Praat: doing phonetics by computer (Version 4.0.4) [Computer program]. Retrieved August 2006, from http://www.praat.org/

Couper-Kuhlen, Elizabeth, and Margret Selting
1996 Towards an interactional perspective on prosody and a prosodic perspective on interaction. In *Prosody in conversation*, Elizabeth Couper-Kuhlen, Margret Selting (eds.), 11-56. Cambridge: Cambridge University Press.

Cruttenden, Alan
1997 *Intonation*. 2nd ed. Cambridge: Cambridge University Press.

Féry, Caroline
1993 *German intonational patterns*. Tübingen: Niemeyer.

Féry, Caroline
2007 Information Structural Notions and the Fallacy of Invariant Correlates. In *Working Papers of the SFB632 Interdisciplinary Studies on Information Structure (ISIS)*, C. Féry, G. Fanselow and M. Krifka (eds.) 6: 161-184. Potsdam: Universitätsverlag Potsdam.

Fletcher, Janet, Esther Grabe, and Paul Warren
2005 Intonational variation in four dialects of English: the high rising tune. In
Prosodic typology. The phonology of intonation and phrasing, Sun-Ah Jun (ed.), 390-409. Oxford: Oxford University Press.

Gilles, Peter
2005 *Regionale Prosodie im Deutschen. Variabilität in der Intonation von Abschluss und Weiterweisung.* Berlin: de Gruyter.

Gilles, Peter, and Jörg Peters (eds.)
2004 *Regional variation in intonation.* Tübingen: Niemeyer.

Grabe, Esther
1998 *Comparative intonational phonology: English and German.* Wageningen: Ponsen and Looijen.

Grabe, Esther, and Brechtje Post
2002 Intonational variation in the British Isles. In *Proceedings of the Speech Prosody 2002 Conference, 11-13 April 2002, Aix-en-Provence: Laboratoire Parole et Langage,* B. Bel, I. Marlin (eds.), 343-346.

Grabe, Esther, Post, Brechtje, Nolan, Francis
2000 Pitch accent realization in four varieties of British English. *Journal of Phonetics* 28: 161-185.

Grice, Martine, and Stefan Baumann
2002 Deutsche Intonation und GToBI. *Linguistische Berichte* 191: 267-298.

Grice, Martine, Ladd, Robert D., and Amalia Arvaniti
2000 On the place of phrase accents in intonational phonology. *Phonology* 17: 143-185.

Gussenhoven, Carlos
1983 Focus, mode and the nucleus. *Journal of Linguistics* 19: 377-417.

Gussenhoven, Carlos
2004 *The phonology of tone and intonation.* Cambridge: Cambridge University Press.

Jacobs, Joachim
1997 I-topikalisierung. *Linguistische Berichte* 168: 91-133.

Kügler, Frank
2007 *The Intonational Phonology of Swabian and Upper Saxon.* Tübingen: Niemeyer.

Ladd, Robert
1996 *Intonational phonology*. Cambridge: Cambridge University Press.

Mehlhorn, Grit
2001 Produktion und Perzeption von Hutkonturen im Deutschen. *Linguistische Arbeitsberichte* 77: 31-57.

Ohala, John
1983 Cross-language use of pitch: an ethological view. *Phonetica* 40: 1-18.

Peters, Jörg
2006 *Intonation deutscher Regionalsprachen*. Berlin: de Gruyter.

Pierrehumbert, Janet
1980 *The phonology and phonetics of English intonation*. Ph.D. thesis, MIT Cambridge.

Pierrehumbert, Janet, and Julia Hirschberg
1990 The meaning of intonational contours in the interpretation of discourse. In *Intentions in communication,* Paul Cohen, J. Morgan, and M. Pollack (eds.), 271-311. Cambridge: MIT Press.

Selting, Margret
2000 The construction of units in conversational talk. *Language in Society* 29: 477-517.

Selting, Margret and Elizabeth Couper-Kuhlen
2000 Argumente fr die Entwicklung einer 'interaktionalen Linguistik'. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion.* (http://www.gespraechsforschung-ozs.de).

Truckenbrodt, Hubert
2005 On intonation phrase boundaries in German. *Linguistische Berichte* 203: 273-296.

Uhmann, Susanne
1991 *Fokusphonologie*. Tübingen: Niemeyer.

A model for the quantification of pitch accent realisation[1]

Frank Kügler

1. Introduction

This paper is concerned with *realisational* differences of intonation, i.e. a different phonetic implementation of an identical phonological category (Ladd, 1996: 199ff). In particular, we are concerned with two aspects of realisational differences, tonal alignment and pitch excursion.[2] The aim of this paper is to propose a two-dimensional model of pitch accent realisation, where both the horizontal (alignment) as well as the vertical level (excursion) of pitch implementation play an important role in characterizing a certain pitch accent category and establishing cross-dialectal differences.

With regards to tonal alignment, several different studies have shown that tones are anchored at certain parts of the segmental string, and that this anchoring is fairly invariant for a given tonal category (cf. section 1.1). We may call this the *invariance hypothesis* of tonal alignment, which predicts that the specific alignment of tones determines which tonal category is involved (e.g. Pierrehumbert and Steele, 1989, for the distinction between English L*H and LH* pitch accents). A given accent category may also differ in its specific alignment across dialects (e.g. Atterer and Ladd, 2004, for L*H accents in Northern and Southern German).

With respect to pitch excursion, very little is known about its particular role as a dialect–and language–specific cue. In a comparison of falling accents Gilles (2005: 164 ff.) establishes that this parameter varies across German dialects. Thus, excursion of accents appear to be a dialect-specific phonetic cue.

The proposal of a two-dimensional pitch accent model that quantifies the variation of individual tonal categories cross-dialectally is based on alignment and excursion measurements carried out for the two German varieties Swabian and Upper Saxon. We assume that the two factors play an important role in capturing intonational variation across dialects.

The structure of this paper is as follows. Sections 1.1 and 1.2 introduce the relevant aspects of tonal alignment and pitch excursion. In section 2 we present the method for obtaining alignment and excursion data as well as the results of the measurements for both dialects. In section 3 the comparison between Swabian and Upper Saxon nuclear rises is presented showing that dialects differ in alignment and excursion. In addition, a comparison of further nuclear contours illustrates how the model works. Section 4 concludes the paper.

1.1. The horizontal level of accent realisation

Accentual timing has been relevant for the theory of intonational phonology from the very beginning. Bruce (1977) has shown that the distinction between Accent I and Accent II words in Swedish is based on a timing difference in word accentual fall, which starts before the accented syllable in Accent I words, but whitin the accented syllable in Accent II words. Accordingly, the timing of the tonal correlate of sentence accent (i.e. focus accent) depends on accent type: the focal high tone aligns earlier for Accent I than for Accent II words (Bruce, 1977: 50).

The alignment of tones has been studied from several different angles and with different aims. First, and probably most importantly from a theoretical point of view, the assumption that tonal targets rather than tonal movements are the relevant categories of intonational phonology has been proven to be true (Silverman and Pierrehumbert, 1990; Arvaniti et al., 1998, 2000; Ladd et al., 1999, 2000; Dilley et al., 2005). This aspect refers to the so-called ‘levels vs. configurations’ debate; for overviews see Bolinger (1951), Ladd (1980: 9ff), Pierrehumbert (1980: 28ff), and Ladd (1996: 59ff). The levels approach of intonation predicts that the turning points of an accent are aligned invariantly at a certain point in the segmental string of an utterance (*invariance hypothesis*). Such an anchor point may, for instance, occur at the left edge of the stressed syllable’s onset or at the beginning of the stressed syllable’s nucleus.

From a phonological point of view the validity of a given phonological category has been tested by means of its phonetic alignment (Pierrehumbert and Steele, 1989; Kohler, 1991b,a; Arvaniti et al., 1998; Ladd and Schepman, 2003). In the 1990s a particular interest in the phonetics of intonation arose from a point of view of speech technology, where synthesized speech sounded unnatural, but was improved by implementing phonetic details of pitch accent realisation (Rietveld and Gussenhoven, 1995; Prieto et al., 1995). Further, there is a growing body of single language descriptions that focuses both on detailed interpretation of tonal categories (e.g. Gilles, 2005; Kügler, 2004) and variation between language varieties (Atterer and Ladd, 2004; Willis, 2003). And finally, the alignment of tones is related to segmental gestures in oder to learn more about synchronous aspects of laryngeal and supra-laryngeal articulatory gestures (Mücke et al., 2006; Baumann et al., 2007).

Across these theoretical interests, different aspects of accents and different effects on the realisation of accents have been studied, such as prenuclear accents (Arvaniti et al., 1998, 2000; Silverman and Pierrehumbert, 1990; Ladd et al., 1999, 2000; Atterer and Ladd, 2004), nuclear accents (Rietveld and Gussenhoven, 1995; Gilles, 2005; Schepman et al., 2006), accentual peaks (Silverman and Pierrehumbert, 1990; House and Wichman, 1996), or lows (Ladd and Schepman, 2003; Dilley et al., 2005), or both (Arvaniti et al., 1998, 2000; Ladd et al., 1999, 2000), and contextual effects such as prosodic structure (Prieto et al., 1995), syllable (Rietveld and Gussenhoven, 1995; Prieto et al., 1995; Ladd et al., 2000), and discourse structure (House and Wichman, 1996; Wichman et al., 2000).

The conclusion of previous studies on tonal alignment is that tonal targets are aligned independently of each other (in particular Dilley et al., 2005) and with respect to segmental landmarks in or around the accented syllable. That is, duration and slope of movement vary dependent on the segmental makeup of stressed syllables. In addition, pitch accents appear to be aligned language specifically. Therefore, anchor points of an identical tonal category differ between languages. Also, starred tones seem to align more stable than leading and/or trailing tones (Prieto 2009).

1.2. The vertical level of accent realisation

In the study of intonational variation, the parameter ‘ pitch excursion’ between two accentual targets has received relatively little attention. The effect of broad and narrow focus on the alignment and pitch excursion of falling accents in Standard German has been investigated (Kügler et al., 2003). It appears that initial pitch excursion is smaller in narrow focus since speakers tend to start higher. Tonal scaling remains the same in different conditions, and depends also on the position of the accent in the phrase.

In a comparison of the intonation in urban German varieties, Gilles (2005) suggests that pitch excursion functions as a dialect specific cue, where the excursion of falling accents appears to be somehow more reliable and less variant than that of rising accents. The absolute excursion size ranges from 6.43 semitones in Dresden to 9.56 semitones in Duisburg (Gilles, 2005: 165). According to Gilles, the main difference between varieties is the excursion and the shape of F0 within the accented syllable. In general, Gilles identifies dialect differences on the basis of excursion measurements, in particular, neighboring dialects have very similar excursions and minor excursions are found in Eastern German dialects while larger excursions are found in Western German dialects. To reduce the effect of speaker specific pitch ranges, Gilles relates the individual measurements to the global pitch range of a speaker.

In intonational phonology, measuring tonal excursion may also function as a diagnostic for a tonal category. Ladd and Schepman (2003) investigated whether a sequence of two high pitch accents in American English consists of two H* pitch accents or a H* L+H* sequence. In varying the number of intervening syllables between the two accentual peaks, Ladd and Schepman showed that in sequences of high pitch accents distinct L and H targets exist, and concluded therefore that English high pitch accents should be analysed as L+H*.

In sum, the aspect of pitch excursion appears to carry crucial information for the analysis of pitch accents. Moreover, pitch excursion appears to be relevant to capture certain intonational variation between dialects, yet more research on pitch excursion is needed to fully understand its role with respect to pitch accent realisation.

2. Accent realisation in two varieties of German

2.1. Introduction

This part of the paper shows that phonological identical contours differ with respect to phonetic implementation across language varieties. Notwithstanding, similarities can also be identified. We will further show that the accent shape on the accented syllable of functionally identical, yet phonologically distinct contours contribute to the dialectal characteristics of accentuation.

Based on the discussion above we assume that accent implementation depends on two parameters: tonal alignment and pitch excursion. Both levels of phonetic implementation appear to carry crucial information for the identity of a tonal category and may therefore be useful for the detection of cross-dialectal differences in accent implementation.

Our model of pitch accent realisation takes these two factors into account. The model is based on a bottom-up approach to model variation in accent realisation.

2.2. Analytical procedures

The analyses and measurements reported here are based on the data in Kügler (2007: 142, 145). The phonological analysis of Swabian is summarized in (1) and that of Upper Saxon in (2). In curly brackets, a tonal grammar displays prenuclear, nuclear pitch accents and boundary tones. Sub- and superscripted numbers for prenuclear pitch accents indicate their possible occurrence from 0 to *n* within an intonation phrase. The L tone in parenthesis indicates a potential extension of a nuclear contour; see Gussenhoven (1984) for a proposal of derived nuclear contours and tonal suffixes, and see Kügler (2006) for a proposal of tonal affixes. OCP refers to the *Obligatory Contour Principle* (Leben, 1980), which prohibits the occurrence of a sequence of two identical tones. A tonal grammar thus reads as an intonation phrase that consists of a possible prenuclear accent, an obligatory nuclear pitch accent, and an optional L tone as a tonal affix. In Swabian any intonation phrase is additionally realized with an obligatory boundary tone, while in Upper Saxon the boundary tone is optional.

(1) A tonal grammar of Swabian (Kügler, 2007: 93)

$$\left\{\begin{array}{c}\text{H*L}\\ \text{L*H}\end{array}\right\}_0^n \left\{\begin{array}{c}\text{L*H}\\ \text{L*HL}\end{array}\right\} (\text{L}) \left\{\begin{array}{c}\text{H\%}\\ \text{L\%}\end{array}\right\}$$

OCP

(2) A tonal grammar of Upper Saxon (Kügler, 2007: 134).

$$\left\{\begin{array}{c}\text{H*L}\\ \text{L*H}\end{array}\right\}_0^n (\text{L}) \left\{\begin{array}{c}\text{H*L}\\ \text{L*H}\end{array}\right\} \quad \{\ \text{H\%}\ \}$$

According to (1) and (2), there is only one contour that is tonally equivalent between Swabian and Upper Saxon: the simple rise. In the data two versions of the simple rise have been found, the simple rise with an additional rise at the intonation phrase boundary (L*H H%), and the simple rise with high plateau (L*H). Both contours have been analysed as being variants of an underlying nuclear simple rise L*H H% (Kügler, 2007). In case of a simple rise, L*H H%, we cannot reliably identify the high trailing tone in all cases. Therefore, the comparison in section (3) concentrates on the simple rise with following high plateau (L*H) to illustrate the differences in phonetic implementation between Swabian and Upper Saxon. We will further contrast a phonetically similar but phonologically different contour, with neutral accentuation in each dialect to show how the two parameters of accent implementation characterise dialect specific features.

The data for the present study come from a map task corpus. Target words were identified that contain a sonorant onset, and in closed syllables, at least one sonorant coda consonant. According to German phonology, syllables with no coda contain a phonologically long vowel (e.g. Féry, 2000; Wiese, 1996), e.g. the second syllable in the word *Arena* [a.ˈʁeː.nɑ] ‘arena’. Closed syllables consist either of a phonologically long or short vowel (ibid.), e.g. *Mehl* [ˈmeːl] ‘flour’ and *Lamm* [ˈlam] ‘lamb’.

The F_0 has been extracted with Praat (Boersma and Weenink, 2007), using a Hanning window of 0.4 seconds length with a default 10 millisecond analysis frame. Obvious errors of the F_0 algorithm (e.g. octave jumps) have been corrected by hand, and the contour has been smoothed using the Praat smoothing algorithm (frequency band 10 Hz) to diminish microprosodic perturbations. All frequency measurements were semi-automatically done using a script that detects the lowest or highest F_0 value within a given domain (for low tone detection, the domain is either the accented or post-accented syllable, for high tone detection, it is the accented word).

2.2.1. *Alignment*

For the alignment of of tones, the segmentation points given in (3) were identified in each intonation phrase (some of these points are identical to the ones in Atterer and Ladd (2004)). Fig. 1 illustrates the labeling.

(3)

C0	the beginning of the onset of the accented syllable
V0	the beginning of the nucleus of the accented syllable
V1	the end of the nucleus (= the beginning of the coda if any)

C1	the end of the syllable rime (end of the accented syllable, i.e. after the last coda consonant, if a word contains more than one)
L1	the first low tonal target (beginning of any rising part of an accent)
H	the F_0 maximum in rising or falling accents
L2	the second low tonal target (the end of any falling part of an accent)
BT	the boundary tone (relevant for excursion measurements below)

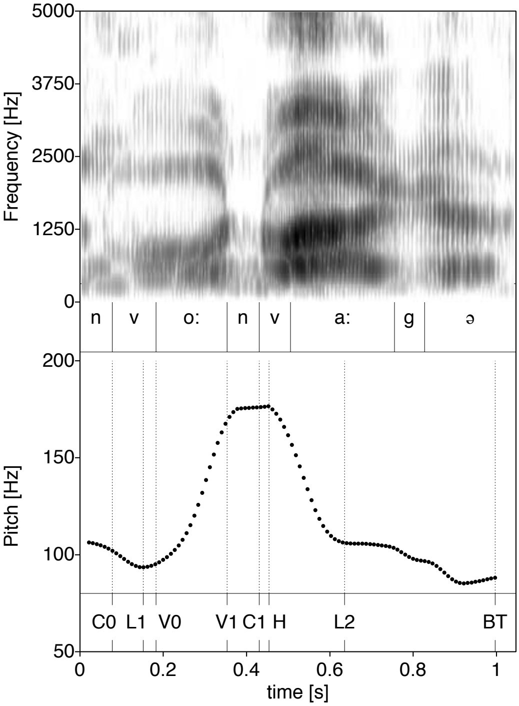


Figure 1: An intonation phrase containing the target word *Wohnwagen* ‘caravan’, showing five segmental boundaries and three F_0 labels given in (3).

Calculation of tonal alignment

Following Atterer and Ladd (2004), we calculated the absolute distance of tonal targets in relation to segmental boundaries. The maximum and minimum of tones (T) were calculated either with reference to V0 or C1, cf. Equations 1 and 2.

$$\mathrm{T}_{\mathrm{V0}_{ref}} = \mathrm{T} - \mathrm{V0} \quad (\mathrm{T} = \text{tone}) \tag{1}$$

$$\mathrm{T}_{\mathrm{C1}_{ref}} = \mathrm{T} - \mathrm{C1} \quad (\mathrm{T} = \text{tone}) \tag{2}$$

With Equation 1, we calculated the alignment of accentual tones (L* and H*) in either dialect; with Equation 2, we calculated the alignment of the trailing tones (H and L) in either dialect. A negative value indicates that the tone is realised before the segmental boundary, while a positive value indicates that the tone is realised after the particular segmental boundary. We are aware of the fact that the exact way of expressing alignment data is debatable (in particular, see appendix A in Atterer and Ladd, 2004), yet we do not attempt to solve that issue here. By using Equation 1, we also provide data that is directly comparable to Atterer and Ladd (2004) and other published data (for example Ladd et al., 1999; Arvaniti et al., 1998).

The alignment data of Swabian and Upper Saxon nuclear rises are reported in Table 1. In addition, the alignment data of the neutral declarative nuclear contours for each dialect are presented as well as the nuclear rise-fall. The table shows the mean alignment of the accentual and trailing tone. The data comprised target sonorant open and closed syllables (see above). The alignment of the accentual tone is given relative to the beginning of the accented syllable's nucleus (V0), and that of the trailing tone (H and L) is given relative to the accented syllable's end (C1). The data are given in milliseconds. For the purpose of contour comparison in the proposed model below, the alignment data of Table 1 are further calculated relative to the syllable's duration.

Table 1: Mean alignment data for Swabian and Upper Saxon nuclear rising contours, nuclear rise-falls and neutral accentuation. The columns show the distance in ms between an F_0 label (L or H) and a segmental boundary (V0 or C1). A negative value indicates that the F_0 label occurs before the segmental label; in case of the rise-falls, only the data of the first two tones are displayed.

	Contour		n	T_{ref}V0	T_{ref}C1
Swabian	L*H	(simple rise)	20	-4.0	-10.4
	L*H+L	(modified rise-fall)	22	-0.6	-36.7
	L*H L%	(neutral rise-fall)	56	1.4	-68.5
Upper Saxon	L*H	(simple rise)	15	3.0	14.5
	L+H*L	(modified simple fall)	37	-14.3	119.9
	H*L	(neutral simple fall)	42	15.7	65.9

From table 1 it can be seen that Swabian and Upper Saxon differ in the realisation of the trailing tone. While in Swabian both accent tones are realised within the accented syllable, in Upper Saxon the trailing tone of the accent aligns on the post-accentual syllable. Peak alignment in Swabian differs significantly between L*H L% and L*H $(df = 46, t = -2.60, p < 0.006)$ contours.[3] For Upper Saxon, alignment of the trailing tone differs; however, since two phonologically distinct tones are involved (L and H) we do not compare these values statistically.

For T*-tone alignment we neither observe a difference between the contours nor between the dialects. On average, an accentual low or high tone is realised near the beginning of the nucleus of the accented syllable. This means that Swabian and Upper Saxon do not differ in L*-tone alignment with regard to the simple rise contour.

To express the alignment of tones in a comparable way, we need to express the actual alignment measurements in terms of the relative distance between the tonal target and a segment or syllable. Based on the reported findings, low tone alignment is best analysed with reference to the beginning of the accented vowel (V0). The relative low tone alignment has yet to be related to the accented syllable in order to compare the data. Therefore, we choose to calculate L alignment by means of the distance between the low tone and the end of the onset consonant (V0) divided by syllable duration (C1 – C0) and multiplied by 100 (cf. Equation 3a).

Alignment of the high tone is best accounted for by calculating the relative distance between the high tone and the end of the accented syllable (C1) divided by the duration of the accented syllable (C1 – C0). The end of the accented syllable appears to be a reasonable segmentation boundary, since the high tone might either be realised in the accented syllable, thus before the end, or in the next syllable, thus after the end (see also Atterer and Ladd, 2004, for this measurement) (cf. Equation 3b).

$$(a)\;\; L(\%) = \frac{(\bar{x}_L + \bar{x}_{onset})}{\bar{x}_{syll}} * 100 \qquad (b)\;\; H(\%) = \frac{(\bar{x}_H + \bar{x}_{onset})}{\bar{x}_{syll}} * 100 \qquad (3)$$

The calculation is based on the means reported in Table 1. We additionally calculated mean onset duration and mean syllable duration for Swabian and Upper Saxon ($\bar{x}_{onset} = 82.5$ ms, $\bar{x}_{syll} = 303.0$ ms, $n = 98$ for Swabian, and $\bar{x}_{onset} = 70.0$ ms, $\bar{x}_{syll} = 272.9$ ms, $n = 94$ for Upper Saxon). An independent samples t-test for onset duration reveals no significant difference between Swabian and Upper Saxon $(p > 0.05)$. Therefore, mean onset duration is based on mean onset duration of Swabian and Upper Saxon onsets $(\bar{x}_{onset} = 76.9$ ms, $n = 192)$. An independent samples t-test for syllable duration reveals a significant difference between Swabian and Upper Saxon accented syllables $(df = 206,\ t = 1.79,\ p < 0.05)$. Therefore, tonal alignment with respect to syllable duration is based on each dialect's mean syllable duration.

The means of Swabian alignment data are $\bar{x}_L = -2.9$ for L-alignment, and $\bar{x}_H = -10.4$ for H alignment (cf. Table 1 above). The data for L* are grouped together in Swabian as reported in Kügler (2007), since we did not find any significant difference for L* alignment across contour type. The means of Upper Saxon L*H are $\bar{x}_L = 1.5$ and $\bar{x}_H = 12.1$ (cf. Table 1 above).

2.2.2. *Excursion*

For pitch excursion, we first have converted the frequency measured in Hz into semitones. We follow Nolan (2003), who has demonstrated that the semitone scale fits best the intonational equivalence (note also Ladd, 1996, 260ff for the notation of semitones with respect to pitch range). The semitone scale also allows for a normalization across gender. A female voice has a higher pitch level than a male voice. This is normalized in the semitone scale since this scale is based on frequency intervals, taking into account that a frequency interval of 50 Hz between 50 and 100 Hz is not equal to an interval between 200 and 250 Hz (see Reetz, 1999, for instance, for technical details of the semitone scale). The conversion into semitones is made according to Equation 4 with a reference of 100 Hz.[4]

$$f(\text{st}) = 12\log_2\left(\frac{f(\text{Hz})}{100\text{ Hz}}\right) \tag{4}$$

Based on semitones, we calculated the pitch excursion by subtracting the low pitch of a pitch accent from the high pitch value (Equation 5a) irrespective of accent type. We concentrated on the rise in rising accents. In order to provide a baseline for the data plot, we calculated the mean low tone in relation to the mean low boundary (BT) (Equation 5b).

$$(a)\ \ E\ [\text{st}] = \text{H}_{\text{st}} - \text{L}_{\text{st}} \qquad (b)\ \ L\ [\text{st}] = \text{L}_{\text{st}} - \text{BT}_{\text{st}} \tag{5}$$

2.3. Results

Table 2 presents the results of the alignment and excursion calculations. The first two columns report the alignment measurements of the low and high tone in relation to the duration of the accented syllable. These data are given as percentages; 0% reflects the beginning of the rime (nucleus) of the accented syllable; a value below 100% means that the tone is realised within the syllable; a value over 100% in the following syllable. The right two columns report the excursion measurements according to Equations 5a and b. The excursion data are given in semitones.

An independent samples t-test for Excursion (E) reveals a significant difference between Swabian and Upper Saxon (E [st] $df = 21$, $t = -1.44$, $p < 0.05$); for the baseline (L tone in relation to the low boundary tone), no significant difference between Swabian and Upper Saxon can be found (L [st] $p > 0.05$). For alignment, we find a significant effect for H alignment ($df = 71$, $t = -1.81$, $p < 0.05$), yet no effect for L* alignment ($p > 0.05$).

Table 2: Data of L and H alignment as percentages of syllable duration, L minimum and Excursion (E) in semitones.

	L [%]	H [%]	L [st]	E [st]
Swabian L*H	24.4	92.0	2.45	7.01
Upper Saxon L*H	28.7	104.4	2.25	8.55

3. Realisational differences of pitch accents – a model

This part of the paper presents the comparison of nuclear accents based on the data of Table 2 above. As can be seen, excursion size differs between the dialects. The lower range is similar between the dialects, Swabian speakers implement a higher rise than Upper Saxon speakers.

Figure 2 illustrates that the dialects differ in some of the parameters. In terms of alignment, the differences are in the realisation of the high tone; in terms of excursion, the differences are in excursion size. In addition to dialect- and language-specific alignment of tones (Atterer and Ladd, 2004), the parameter ‘ pitch excursion’ appears to signal dialect specific intonational features as well.

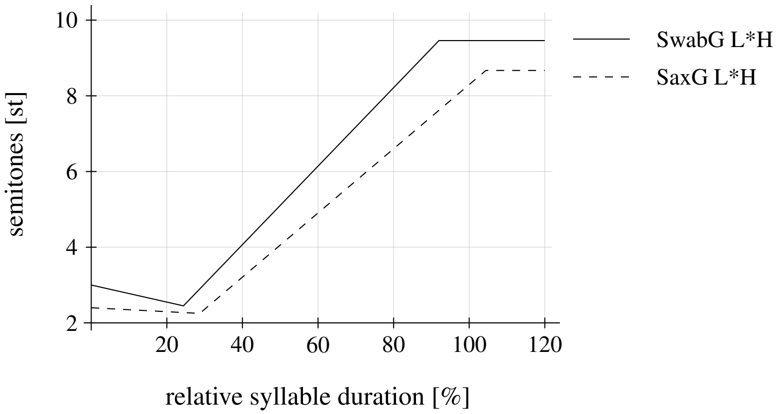


Figure 2: A comparison of alignment and excursion between Swabian and Upper Saxon simple rising accents in the simple rise contour (L*H).

For L-alignment, we find no significant difference between Swabian L* and Upper Saxon L*, meaning that the position of the accentual low tone relative to the onset and syllable duration is identical in the two dialects. Comparing this result to Atterer and Ladd (2004) who have found a significant difference between Northern and Southern (Bavarian) German, we can conclude that for rising accents – a L*H nuclear contour – there is no difference between South-Western and Central-Eastern German. Recall that Atterer and Ladd (2004) analyse prenuclear rising accents while we analyse nuclear ones. The similar implementation of accentual low tones might be an indicator of a perceived similarity between Swabian and Upper Saxon (many of our informants have reported that these dialects do not differ that much intonationally).

Our model contains two factors that influence pitch accent realisation. Gilles (2005) proposed a slightly different model to address these factors. Gilles refers to global F_0 range of a speaker and relates the individual accent realisations to this range. For rising contours, Gilles (2005: 270) observes no regional difference. Further, he calculates the exact contour shape of the accented and postaccented syllable. This procedure is similar to our model, but differs in two aspects. First, Gilles expresses tonal alignment in terms of distance relations of the tone in relation to the duration of the accented syllable, and not in relation to onset-nucleus-boundary as we do and which has been proposed by Atterer and Ladd (2004). Second, the excursion is expressed in terms of the relative excursion of the accent and not in terms of semitones.

Apart from the fact that the global range measurements reported in Gilles (2005) cannot be assumed to be a clear dialect-specific cue, we would not rely on this kind of measurement for others reasons. Given the observation that one can identify a speaker's dialect on the basis of a single utterance, it cannot be based on a calculation procedure that relies on global means of pitch range. We believe that our calculation method reveals obvious differences between the dialects, though we are aware of the fact that different measurements might show even better results, e.g. different normalization calculations or alignment calculations. Moreover, we believe that a calculation which is based on the accented syllable reflects the aspects of accent realisation more properly since it is the syllable which is the domain of tonal implementation–at least for intonation languages.

There might be a further objection against a global pitch range: the question of defining the relevant measuring points. Gilles (2005: 85) does not explicitly discuss the choice of measuring points, i.e. which highs and lows reflect global pitch range. Möhler and Mayer (1999) propose a calculation of pitch range based on pitch accents and boundary tones. A different proposal excludes the first and last pitch accent for the calculation of pitch range (Patterson, 2000). Given the fact that it has not been established which method best models excursion global pitch range and its relation to pitch accent implementation remains an open issue.

If we apply our model to illustrate the F_0 contour on the accented syllable for functionally identical, yet phonologically distinct, contours we identify crucial differences in the phonetics of accented syllables across dialects. Given that the accented syllable is the perceptually salient and the most prominent part of the nuclear contour, the differences in phonetics may reveal the relevant cues on which speakers base their perception on. Consider for instance the modified Swabian rise-fall (L*H+L) and the modified Upper Saxon simple fall (L+H*L). Functionally, these contours share their expression of narrow focus, and they convey a meaning of a 'significant ADDITION' of information to the interlocutors' conversational background (Kügler, 2007). The shape of the respective accent is illustrated in Figure 3, where differences in tonal alignment and excursion sizes become obvious.

The contours on the accented syllable shown in Figure 3 are similar, but not identical. Despite the difference in phonological form, there are subtle differences in the phonetic realisation of the accented syllable with respect to the alignment of the L and H tone and pitch excursion. The Upper Saxon

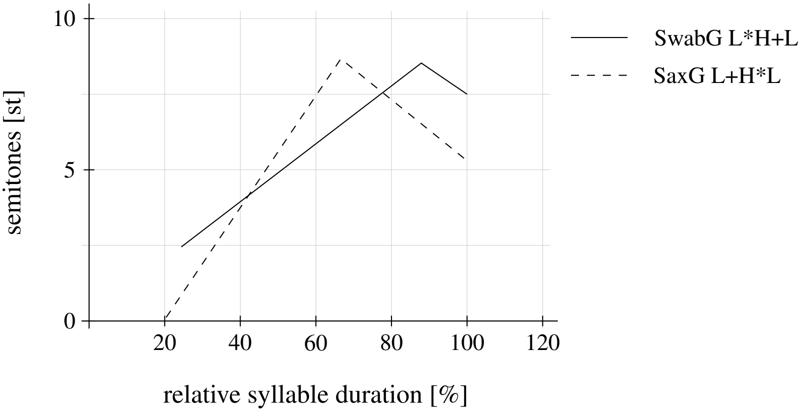


Figure 3: A comparison of alignment and excursion between in rise-fall contours, Swabian L*H+L, Upper Saxon L+H*L.

accent shape is realised earlier than the Swabian one. In addition, the Upper Saxon accent shape has a larger excursion than the Swabian one.

An even more distinct picture arises when we compare the case of neutral accentuation between Swabian and Upper Saxon; According to the analysis in (Kügler, 2007) Swabian has a rise-fall (L*H L%) while Upper Saxon a simple fall (H*L) . The difference of accent shape on the accented syllable is shown in Figure 4.

For Swabian, the accent shape of the accented syllable is rising in neutral accentuation, for Upper Saxon it is falling. This observation may be seen already in our phonological analysis, where we analyse a rising pitch accent for Swabian (L*H) and a falling one for Upper Saxon (H*L). Thus, one may argue against this comparison that differences in phonological form predict differences in phonetic form. This objection does not hold, however, if we compare the Swabian and Upper Saxon data with neutral accentuation in Standard German, which is represented by the dotted line in Figure 4. According to Grabe (1998a) and Féry (1993), in neutral accentuation Standard German has a simple falling accent (H*L), which is phonologically similar to the Upper Saxon simple fall. The phonetic form, however, differs from Upper Saxon since Grabe has shown that falling accents align their accentual high tone at the right edge of the syllable rime. In Figure 4, we have plotted Grabe's alignment data (without reference to excursion since no data are available) against the Swabian and Upper Saxon realisation.[5]

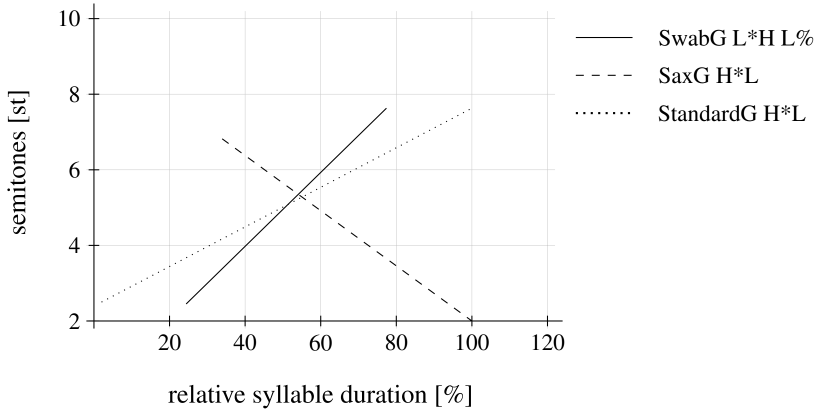


Figure 4: A comparison of alignment and excursion between Swabian and Upper Saxon simple rising accents in the simple rise contour (L*H).

Comparing all three accents in Figure 4, the Standard German simple fall is fairly similar in phonetic form on the accented syllable to the Swabian rise-fall; note that these two accents differ in phonological form. A phonologically identical accent, on the contrary, appears in a different accent shape on the accented syllable, cf. the Upper Saxon simple fall compared to the Standard German one.[6] Thus, the phonetics of the accent shape on the accented syllable is a crucial property of a given language variety irrespective of the phonological form of the contour.

An interesting question that we leave for future research is the perceptual relevance of these aspects of accent realisation. One wants to investigate how much the accent shape of the accented syllable influences the perception of a phonological tone. The acoustic measurements at least appear to point to dialect specificity when it comes to the implementation of accents on the accented syllable.

4. Summary

This paper is concerned with phonetic aspects of tonal implementation, in particular with tonal alignment and pitch excursion. In a comparison between two German dialects, Swabian and Upper Saxon, we identify two parameters that, taken together, allow quantification of intonational variation with respect to tonal implementation across language varieties. Our model combines tonal alignment with respect to the accented syllable and pitch excursion for implementing pitch accents. Comparing a phonologically similar contour, L*H, shows that certain phonetic aspects of pitch accents differ across dialects. We observe that H-tone alignment as well as excursion size differ significantly, while L-tone alignment and the scaling of low tones is similar between the two dialects.

The factors of tonal alignment and tonal excursion have been shown to play a crucial role in the characterization of dialect specific accent implementation. In addition, we have presented evidence that phonologically different but functionally similar contours differ in their exact phonetic accent implementation. The rise-fall contours in both dialects seem to be similar phonetically. The detailed phonetic measurements of alignment and excursion however show subtle differences between the two dialects. The interesting question whether these subtle differences in the acoustics of pitch accents would hold perceptually must be left open for future research.

Notes

1. This work has been part of the author's PhD project on intonational variation in German dialects published as Kügler (2007). Special thanks go to Caroline Féry and Jörg Mayer for their advice and support during the writing of the dissertation, to an anonymous reviewer, and to Peter Gilles and Ruben van de Vijver who thoroughly commented on an earlier version of this paper. The usual disclaimers apply.
2. A further realisational aspect concerns the implementation of pitch under time pressure–truncation or compression–which we leave aside here in order to concentrate on the realisation of pitch accents on fully voiced words (see Grabe, 1998a,b; Grabe et al., 2000; Kügler, 2003, 2004; Gilles, 2005; Peters, 1999, for analyses of pitch under time pressure).
3. Peter Gilles raised the question of normalizing across speakers to accommodate for differences in speaking tempo. Given the fact that segmental anchoring remains constant under changes of speech rate (Ladd et al., 1999) we assume that the alignment measurements were certainly not affected by speaking rate.
4. Male voices may reach negative semitones because of the reference of 100 Hz. Since the absolute tonal values are reported in relation to the low boundary tone within the phrase, negative values do not occur. A low pitch accent is always higher than a boundary tone within one phrase, see Liberman and Pierrehumbert (1984) for the fact of constant final lows in intonation. The plots shown in Figures 2 to 4 only present a positive scale.

5. As Peter Gilles points out, the excursion of the Standard German rise within the accentual syllable cannot be taken as fact since Grabe (1998a) does not provide excursion data. However, according Grabe a rise occurs in the accented syllable and aligns as shown in Figure 4. The actual rise starting at around 2 st is only for illustrative purposes.
6. Note that the simple fall in our data, i.e. recordings from speakers of the urban vernacular of Leipzig appears to differ from speakers of Dresden which belongs to the same dialect area (Gilles, 2005).

References

Arvaniti, A., D. R. Ladd, and I. Mennen
1998 Stability of tonal alignment: The case of Greek prenuclear accents. *Journal of Phonetics* 26: 3-25.

Arvaniti, A., D. R. Ladd, and I. Mennen
2000 What is a starred tone? Evidence from Greek. In *Papers in Laboratory Phonology V*, M. Broe and J. Pierrehumbert (eds.), 119–131. Cambridge: Cambridge Univeristy Press.

Atterer, M. and D. R. Ladd
2004 On the phonetics and phonology of "segmental anchoring" of F0: Evidence from German. *Journal of Phonetics* 32: 177–197.

Baumann, S., J. Becker, M. Grice, and D. Mücke
2007 Tonal and Articulatory Marking of Focus in German. In *Proceedings of the 16th ICPhS*, J. Trouvain and W. J. Barry (eds.), 1029–1032.

Boersma, P. and D. Weenink
2007 Praat: doing phonetics by computer (Version 4.5) [Computer program]. Retrieved from http://www.praat.org/.

Bolinger, D. L.
1951 Intonation: levels versus configurations. *Word* 7: 199-210.

Bruce, G.
1977 *Swedish Word Accents in Sentence Perspective*. Travaux de L'institut de Linguistique de Lund, XII.

Dilley, L., D. R. Ladd, and A. Schepman
2005 Alignment of L and H in bitonal pitch accents: testing two hypotheses. *Journal of Phonetics* 33(1): 115-119.

Féry, C.
1993 *German Intonational Patterns*. Tübingen: Niemeyer.

Féry, C.
2000 *Phonologie des Deutschen. Eine optimalitätstheoretische Einführung*. [LIP 7 and 11]. Postdam: Institut für Linguistik.

Gilles, P.
2005 *Regionale Prosodie im Deutschen. Variabilität in der Intonation von Abschluss und Weiterweisung*. Berlin/New York: de Gruyter.

Grabe, E.
1998a *Comparative Intonational Phonology. English and German*. Ph.D. thesis, Nijmegen University.

Grabe, E.
1998b Pitch accent realization in English and German. *Journal of Phonetics* 26(3): 129-143.

Grabe, E., B. Post, F. Nolan, and K. Farrar
2000 Pitch accent realization in four varieties of Britsh English. *Journal of Phonetics* 28(2): 161-185.

Gussenhoven, C.
1984 *On the Grammar and Semantics of Sentence Accents*. Dordrecht: Foris.

House, J. and A. Wichman
1996 Investigating peak timing in naturally-occurring speech: From segmental constraints to discourse structure. Speech Hearing and Language: Work in Progress 9, Department of Phonetics and Linguistics, University College London.
[http://www.phon.ucl.ac.uk/home/shl9/contents.htm].

Kohler, K. J. (ed.)
1991a *Studies in German Intonation*. Volume 25 of *Arbeitsberichte (AIPUK)*. Kiel: Institut für Phonetik und digitale Sprachverarbeitung.

Kohler, K. J.
1991b Terminal Intonation Patterns in Single-Accent Utterances of German: Phonetics, Phonology and Semantics. In Kohler (1991a), 115–185.

Kügler, F.
2003 The Influence of Vowel Quality and Syllable Structure on the Temporal Organization of Pitch Accents in Berlin German. In *Proceedings of the XVth International Congress of Phonetic Sciences*, 1221-1224. Barcelona.

Kügler, F.
2004 The phonology and phonetics of nuclear rises in Swabian German. In *Regional Variation in Intonation*, P. Gilles and J. Peters (eds.), 75–98. Tübingen: Niemeyer.

Kügler, F.
2006 L-affixation: Evidence from German dialects. In *Proceedings of the third International Conference on Speech Prosody*, R. Hoffmann and H. Mixdorff, eds., 201–204. Dresden: TUDpress Verlag der Wissenschaften.

Kügler, F.
2007 *The Intonational Phonology of Swabian and Upper Saxon*. Tübingen: Niemeyer. [Linguistische Arbeiten 515].

Kügler, F., C. Féry, and R. van de Vijver
2003 Pitch Accent Realization in German. In *Proceedings of the XVth International Congress of Phonetic Sciences*, 1261–1264. Barcelona.

Ladd, D. R.
1980 *The structure of intonational meaning: Evidence from English*. Bloomington: Indiana University Press.

Ladd, D. R.
1996 *Intonational Phonology*. Cambridge: CUP.

Ladd, D. R., D. Faulkner, H. Faulkner, and A. Schepman
1999 Constant segmental anchoring of F0 movements under changes in speech rate. *Journal of the Acoustical Society of America* 106(3): 1543–1554.

Ladd, D. R., I. Mennen, and A. Schepman
2000 Phonological conditioning of peak alignment in rising pitch accents in Dutch. *Journal of the Acoustical Society of America* 107(5): 2685–2696.

Ladd, D. R. and A. Schepman
2003 "Sagging transitions" between high pitch accents in English: experimental evidence. *Journal of Phonetics* 31(1): 81–112.

Leben, R. W.
1980 *Suprasegmental Phonology*. New York: Garland.

Liberman, M. and J. Pierrehumbert
1984 Intonational invariance under changes in pitch range and length. In *Language Sound Structure: Studies in Phonology presented to Morris Halle*, 157–233. Cambridge: MIT Press.

Mücke, D., M. Grice, J. Becker, A. Hermes, and S. Baumann
2006 Articulatory and Acoustic Correlates of Prenuclear and Nuclear Accents. In *Proceedings of Speech Prosody 2006, Dresden, Germany*, R. Hoffmann and H. Mixdorff, (eds.) 297–300. Dresden: TUDpress.

Möhler, J. and J. Mayer
1999 A method for the analysis of prosodic registers. In *Proceedings of the 6^{th} European Conference on Speech Communication and Technology*, volume 2, 735–738. Budapest.

Nolan, F.
2003 Intonational equivalence: An experimental evaluation of pitch scales. In *Proceedings of the XVth International Congress of Phonetic Sciences*, 771–774. Barcelona.

Patterson, D.

2000 *A Linguistic Approach to Pitch Range Modelling*. Ph.D. thesis, University of Edinburgh.

Peters, J.

1999 Trunkierung und Kompression bei finalen Akzentsilben. Vorläufige Resultate einer Untersuchung zum Berlinerischen und Hamburgischen. MS, Potsdam University.

Pierrehumbert, J.

1980 *The phonology and phonetics of English intonation*. Ph.D. thesis, MIT. Distributed by Indiana University Linguistics Club, Bloomington.

Pierrehumbert, J. and S. A. Steele

1989 Categories of Tonal Alignment in English. *Phonetica* 46: 181–196.

Prieto P.

2009 Tonal alignment patterns in Catalan nuclear falls. *Lingua* 119: 865-880.

Prieto, P., J. van Santen, and J. Hirschberg

1995 Tonal alignment patterns in Spanish. *Journal of Phonetics* 23: 429–451.

Reetz, H.

1999 *Artikulatorische und akustische Phonetik*. Trier: Wissenschaftlicher Verlag Trier.

Rietveld, T. and C. Gussenhoven

1995 Aligning pitch targets in speech synthesis: effects of syllable structure. *Journal of Phonetics* 23: 375–385.

Schepman, A., R. Lickley, and D. R. Ladd

2006 Effects of vowel length and “right context” on the alignment of Dutch nuclear accents. *Journal of Phonetics* 34: 1–28.

Silverman, K. E. A. and J. Pierrehumbert

1990 The timing of prenuclear high pitch accents in English. In *Papers in Laboratory Phonology I. Between the Grammar and Physics of Speech*, J. Kingston and M. E. Beckman (eds.), 72–106. Cambridge: Cambridge University Press.

Wichman, A., J. House, and T. Rietveld

2000 Discourse constraints on F0 Peak Timing in English. In *Intonation: Analysis, Modelling and Technology*, A. Botinis (ed.), 163–184. Dordrecht: Kluwer.

Wiese, R.

1996 *The phonology of German*. Oxford: Clarendon Press.

Willis, E. W.

2003 Prenuclear Low Tone Alignment in Dominican Spanish. In *Proceedings of the XVth International Congress of Phonetic Sciences*, 2941–2944. Barcelona.

List of authors

Pia Bergmann
Deutsches Seminar I, University of Freiburg, Germany

Anne Christophe
Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS, Paris, France

Thaïs Cristófaro-Silva
Department of Linguistics, Federal University of Minas Gerais, Brazil

Isabelle Darcy
Department of Second Language Studies, Indiana University Bloomington, USA

Emmanuel Dupoux
Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS, Paris, France

Caroline Féry
Department of Linguistics, Potsdam University, Germany

Robert Allen Fox
Department of Speech and Hearing Science, The Ohio State University, USA

Christoph Gabriel
Department of Romance languages, University of Hamburg, Germany

Daniela Oliveira Guimarães
Department of Linguistics, Federal University of Minas Gerais, Brazil

Silke Hamann
Department of Linguistics, Heinrich-Heine University Düsseldorf, Germany

Frans Hinskens
Department of Variation Linguistics, Meertens Instituut (KNAW),
Vrije Universiteit, Amsterdam, The Netherlands

Barbara Höhle
Department of Linguistics, Potsdam University, Germany

Ewa Jacewicz
Department of Speech and Hearing Science, The Ohio State University, USA

Katherine Kinzler
Department of Psychology, University of Chicago, USA

Alexei Kochetov
Department of Linguistics, University of Toronto, Canada

Frank Kügler
Department of Linguistics, Potsdam University, Germany

Judith Meinschaefer
Department of Romance languages, University of Würzburg, Germany

Trudel Meisenburg
Department of Romance languages, University of Osnabrück, Germany

Jeff Mielke
Department of Linguistics, University of Ottawa, Canada

Marc van Oostendorp
Department of Variation Linguistics, Meertens Instituut (KNAW),
The Netherlands

Susan Ott
Department of Linguistics, Potsdam University, Germany

Franck Ramus
Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS, Paris, France, Institute of Cognitive Neuroscience, University College London, UK

Joseph Salmons
Department of German, University of Wisconsin-Madison, USA

Šárka Šimáčková
Department of English and American Studies, Palacký University Olomouc, Czech Republic

Ruben van de Vijver
Department of Linguistics, Potsdam University, Germany

Index

A

- abstractness 351
- accent 194
- adult language 247, 249, 257, 261
- age group 332
- allomorphy 230
- allophones 194
- allophony 201
- ambisyllabicity 319
- Aoma 27
- approximant 71f., 74, 76f., 87, 89, 92f.
- articulatory constraint
  - degrees of 67
- articulatory gestures 43, 45, 65
- assimilation 44, 145, 265f., 269 – 278, 284ff., 292, 294ff., 353f., 359 – 364, 371
  - nasal 187
  - place 278, 284 – 288, 291, 293, 295, 297
  - voicing 265, 275, 277, 284 – 287, 291, 293 – 297

C

- categorical phonology 185
- chain shifting 99, 101, 120f.
- Cheremis
  - Eastern 24f.
- child language 247f., 257, 260f.
- co-phonologies 186
- compensation 265, 268 – 276, 280, 283 – 286, 290 – 297
  - lexical 268ff., 295, 297
  - phonetic 268, 270, 295
- consonant harmony 257
- consonants
  - palatalized 45, 66
- creaky voice 167
- Czech 128ff., 134ff.

D

- dark *l* 125 – 130, 135f.
- derivation 351, 355, 368
- dialect 353, 357, 359, 362, 365ff., 371f., 377f., 405f., 408, 410, 412 – 417, 419ff.
  - convergence 311
  - divergence 311
  - levelling 312ff., 332, 338, 341
- duration 76 – 85, 87 – 90, 92, 94
- Dutch 259ff., 273f.
  - Antwerp 356ff., 370
  - Brabant 363
  - Hellendoorn 353f., 359f., 362f., 366
  - Limburg 312, 314, 321, 341
  - Ripuarian 321
  - Wilsum 366

E

- Edo 27
- Edoid language 26ff.
- Ehueun 27
- EMMA 47
- English 252, 256, 260f.
  - American 99, 101ff., 117, 265, 277, 286, 288f., 292 – 295
    - Ohio 102
    - Wisconsin 102, 118
- Epie 27
- Eruwai 27

F

- Factor-driven gradient model 203
- faithfulness 351 – 355, 357, 361, 366f., 369ff.
- Featurally Underspecified Lexicon 269
- feature
  - coronal 248, 250, 253

  - dorsal . . . . 186 – 189, 191 – 194, 197f., 200, 206 – 209, 248, 251f., 254f., 257ff., 261
  - labial 253, 258
  - natural class . 18, 21, 25f., 29, 31ff., 36f.
  - place of articulation 248, 257
  - unnatural class . . 26, 29, 31, 33, 36

- Feature Cue Parsing 271
- feature geometry 312
- final devoicing 187
- formant pattern 99, 106, 111
- French167, 171, 173, 179f., 277f., 280f., 284 – 289, 292 – 295, 297f.
  - élision 163
  - enchaînement . . . . 171f.
  - h aspiré . . . . 169 – 174, 179, 181
  - liaison . . . . 171f.
  - Verlan 165
- frequency effects . . . . 141, 152 – 155
- fricative . . 74, 76f., 79f., 86, 88, 91, 148, 151
- Full Transfer Theory 86
- Functional Phonology 200

G

- g deletion 187
- German . . . . 247 – 252, 260f., 273f., 298
  - Cologne . 377f., 380, 382, 384, 386, 388 – 392, 394, 397ff.
  - Northern . . . 188f., 197f., 203, 207f.
  - Standard 207, 407, 418f., 421
  - Swabian . . . . 418f.
  - Swiss 17, 26
  - Upper Saxon . . . . 405f., 409f., 412 – 419
- gestural
  - blending . . . . 46f., 53 – 57, 59, 66f.
  - invariance 46f., 50, 53
  - reduction 47, 56, 59ff., 66
- Gestural Linguistic Model . . . . 43f., 66f.
- glide 94
- glottal stop . . . . . . 167, 170 – 173, 179ff.
- Gradual Learning Algorithm 198

H

- hat pattern . . . . . . 378, 380, 392ff., 398f.
- hiatus . . . . 167f., 173f.
- Hungarian 273, 296

I

- inflection 353, 364
- intonation . . . . 377ff., 382, 386, 388 – 391, 394, 398, 400f., 405f., 408ff., 417, 420
- Italian 215
  - stress 233
  - Tuscan 216

J

- Japanese 274

L

- lambdacisms 126
- lateral liquid . . . . 125, 129f.
- lenition . . . 24, 37, 141, 144ff., 148 – 155
- lexical diffusion 141f., 148, 152
- Lexicon Optimization 258
- linguistic change
  - directionality 337
  - effective causes 324
  - enabling causes 321
  - external motivation 313
  - internal motivation 315

M

- map task 410
- multiple inputs 235
- multiple readout model . . . . 296f.

N

- neutralization 71
- noisiness . . . . 83f.

O

- Ojibwa 362f., 366

Okpe 27
onset . . . . 167, 170 – 174, 179
  cluster . . . . 259ff.
  complex . . . . 247 – 250, 253 – 256, 258 – 261
  singleton 247
opacity 351, 353ff., 357 – 360, 363, 366, 368, 370ff.
opaque process 200
ordered rules 187
OT-CC 198

P

palatalization . . . . 141 – 144
perception grammar . . . 71 – 74, 80, 83f., 86 – 89, 91f.
phonological phrase 225
pitch
  accent implementation 417
  alignment 405, 407
  excursion . . . 405, 407f., 414f., 417, 419f.
Portuguese
  Brazilian . . . . 141f.
prominence
  prosodic . . . 103 – 107, 109, 111ff., 117 – 121
prosodic boundary 192
prosodic constituency 344
prosodic word 224

R

recoverability 351, 359, 362, 369
rhotacism 127
Russian . 26, 29f., 35, 37, 43 – 47, 56, 64

S

schwa . 166, 168, 170f., 174, 176, 179ff., 353, 359f., 362, 366, 370
sound change 99, 102, 117f., 121, 141f., 148, 152
SPE . . 2, 4f., 8, 21, 25, 29, 81, 215, 320, 355
spontaneous speech 378f., 399
standard OT 188
Stochastic OT 199
syllable structure 259f., 318
syntactic expansion 395
syntax-phonology interface 223

T

troncamento 215
typology 249, 256f., 261

U

Ukue 27
Unified Feature Theory . . . . 21, 24f.
Urhobo 27, 37
Uvbie 27

V

variable transformational rules 186
velarisation . . . . 356f.
vowel
  acoustics . . 102, 104, 106, 111, 114
  dynamics 99
  perception . . . . 112, 118ff.

W

word recognition275, 277, 284, 293, 296