Current Issues in Linguistic Theory

The Phonological Spectrum

Volume II: Suprasegmental structure

EDITED BY
Jeroen van de Weijer
Vincent J. van Heuven
Harry van der Hulst

THE PHONOLOGICAL SPECTRUM II

AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE

General Editor
E. F. KONRAD KOERNER
(University of Cologne)

Series IV – CURRENT ISSUES IN LINGUISTIC THEORY

Volume 234

Jeroen van de Weijer, Vincent J. van Heuven and
Harry van der Hulst (eds.)

The Phonological Spectrum
Volume II: Suprasegmental structure

THE PHONOLOGICAL SPECTRUM

VOLUME II: SUPRASEGMENTAL STRUCTURE

Edited by

JEROEN VAN DE WEIJER
Universiteit Leiden

VINCENT J. VAN HEUVEN
Universiteit Leiden

HARRY VAN DER HULST
University of Connecticut, Storrs

JOHN BENJAMINS PUBLISHING COMPANY
AMSTERDAM/PHILADELPHIA

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences — Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

The Phonological Spectrum / Edited by Jeroen van de Weijer, Vincent J. van Heuven and Harry van der Hulst.

p. cm. -- (Amsterdam studies in the theory and history of linguistic science. Series IV, Current issues in linguistic theory, ISSN 0304-0763 ; v. 233-234)

vol. 1. Segmental structure -- vol. 2 Suprasegmental structure.

Includes bibliographical references and index.

1. Grammar, Comparative and general. I. Weijer, Jeroen van de. II. Heuven, Vincent J. van. III Hulst, Harry van der. IV Series.

P217 .P4827 2003

414--dc21 2002038462

Vol. 1 ISBN 1 58811 351 5 (US) / 90 272 4744 7 (EUR.) (Hb; alk. paper)

Vol. 2 ISBN 1 58811 352 3 (US) / 90 272 4745 5 (EUR.) (Hb; alk. paper)

Set ISBN 1 58811 353 1 (US) / 90 272 4746 3 (EUR.) (Hb; alk. paper)

John Benjamins Publishing Co. • P.O.Box 36224 • 1020 ME Amsterdam • The Netherlands

John Benjamins North America • P.O.Box 27519 • Philadelphia PA 19118-0519 • USA

Table of contents

Preface **VII**

Syllables, feet and higher up **IX**

Syllabic structure

Phonetic evidence for phonological structure in syllabification **3**
Heather Goad and Kathleen Brannen

The phonology-phonetics interface and Syllabic Theory **31**
Shosuke Haraguchi

Hungarian as a strict CV language **59**
Krisztina Polgárdi

Syllable structure at different levels in the speech production process: Evidence from aphasia **81**
Dirk-Bart den Ouden and Roelien Bastiaanse

Metrical structure

Quantity-sensitivity of syllabic trochees revisited: The case of dialectal gemination in Finnish **111**
Heli Harrikari

Ternarity is Prosodic Word binarity **127**
Jay I. Rifkin

The status of word stress in Indonesian **151**
Ellen van Zanten, Rob Goedemans, and Jos Pacilly

Prosodic structure

Perceived prominence and the metrical-prosodic structure of Dutch sentences **179**
Karijn Helsloot and Barbertje M. Streefkerk

Phonetic variation or phonological difference? The case of the early versus the late-accent lending fall in Dutch **201**
Johanneke Caspers

On the categorical nature of intonational contrasts, an experiment on boundary tones in Dutch **225**
Bert Remijsen and Vincent J. van Heuven

Author index **247**

Language index **255**

Subject index **257**

Table of contents, volume I **261**

Preface

Vincent J. van Heuven, Harry van der Hulst, and
Jeroen van de Weijer
Leiden University / University of Connecticut, Storrs / Netherlands
Institute for Advanced Sciences, Wassenaar

The two volumes of the *Phonological Spectrum* aim at giving a comprehensive overview of current developments in phonological theory, by providing a number of papers in different areas of current theorizing which reflect on particular problems from different angles. Volume I is concerned with segmental structure, and focuses on nasality, voicing and other laryngeal features, as well as segmental timing. With respect to nasality, questions such as the phonetic underpinning of a distinctive feature [nasal] and the treatment of nasal harmony are treated. As for voicing, the behaviour of voicing assimilation in Dutch is covered while its application in German is examined with an eye to its implications for the stratification of the German lexicon. In the final section of volume I, the structure of diphthongs is examined, as well as the treatment of lenition and the relation between phonetic and phonological specification in sign language.

Volume II deals with phonological structure above the segmental level, in particular with syllable structure, metrical structure and sentence-level prosodic structure. Different syllable structure theories, as well as possible relations between segment structure and syllabic structure, and evidence from language acquisition and aphasia are examined in Section 1. Metrical structure is examined in papers on foot structure, and, experimentally, on word stress in Indonesian. Finally in this volume, there are three laboratory-phonological reports on the intonation of Dutch.

December 2002

Syllables, feet and higher up

Vincent J. van Heuven, Harry van der Hulst, and
Jeroen van de Weijer

This volume deals with phonological structure above the segmental level, in particular with syllable structure, metrical structure and sentence level prosodic structure.

In Section 1, different syllable structure theories, as well as possible relations between segment structure and syllabic structure, and evidence from aphasia are examined. The section starts with an article by Heather Goad and Kathleen Brannen ('Phonetic evidence for phonological structure in syllabification'), which looks at the status of syllable structure from the perspective of language acquisition, presenting evidence for the position that word-final consonants are syllabified as onsets, not codas, in the first stages. Dirk-Bart den Ouden and Roelien Bastiaanse ('Syllable structure at different levels in the speech production process: Evidence from aphasia') looks at the phonetics and phonology of syllable structure in aphasic patients, and shows that phonological structures play a role in the early planning of utterances. The article in this section by Shosuke Haraguchi shows that the order of segments need not be stipulated in the phonological representation for a number of lexical items in Japanese and English, given a sufficiently richly articulated theory of syllable structure. Finally, Krisztina Polgárdi ('Hungarian as a strict CV language') explores the abstractness of underlying representations in arguing that syllable structure in Hungarian is strictly CV underlyingly, a necessary move to explain surface generalizations.

Metrical structure is examined in papers on foot structure, and, experimentally, on word stress in Indonesian. First, Heli Harrikari ('Quantity-sensitivity of syllabic trochees revisited: The case of dialectal gemination in Finnish') examines the claim that syllabic troches are typically quantity-insensitive against the background of segmental processes in various dialects of Finnish. Also with regard to stress, Ellen van Zanten, Rob Goedemans and

Jos Pacilly ('The status of word stress in Indonesian') re-examine the location of primary stress in Indonesian from a laboratory-phonology perspective, and find that this is phrasal rather than lexical in nature. Ternary stress systems, a long-standing problem to phonological theories embracing (strict) binarity, are the topic of Jay Rifkin's 'Ternarity is Prosodic Word binarity', in which he shows that such structures should be analysed as binary-branching prosodic words, not single feet.

Finally in this volume, there are three laboratory-phonological reports on the intonation of Dutch. Karijn Helsloot and Barbertje Streefkerk ('Perceived prominence and the metrical-prosodic structure of Dutch sentences') present a prosodic analysis of Dutch sentences, arguing in favour of a four-level prominence grid and a hierarchy that incorporates both input and output constraints. Johanneke Caspers ('Phonetic variation or phonological difference: The early versus the late-accent lending fall in Dutch') raises the question when a particular melodic shape in the intonation of a language (in this case, Dutch) should be considered to be phonological and when it should be considered as a case of phonetic variation, trying to differentiate between different models used to describe intonational autosegments. Roughly the same problem, the categorization of boundary tones in Dutch, but exploring a different methodology, is tackled by Bert Remijsen and Vincent van Heuven ('On the category of intonational contrasts – an experiment on boundary tones in Dutch').

Syllabic structure

Phonetic evidence for phonological structure in syllabification*

Heather Goad and Kathleen Brannen
McGill University

1. Preliminaries

It has been widely observed that codas are marked vis-à-vis onsets. In adult languages, coda markedness manifests itself in two ways: some languages do not tolerate codas at all (e.g. Cayuvava (Key 1961), Senufo (Clements & Keyser 1983)), while virtually all languages that do permit codas place restrictions on what this position can license. Coda markedness has consequences for the syllabification of consonants at the right edge of words as well. While one view is that final consonants are always syllabified as codas, there are many languages for which this analysis is not motivated. In the 1980s, final consonants were commonly designated as extraprosodic in these types of languages (Steriade 1982; Borowsky 1986; Itô 1986). More recently, it has been observed that there is a striking parallel between right-edge extraprosodic consonants and onsets in many languages; accordingly, extraprosodic consonants have been interpreted as onsets of empty-headed syllables by a number of scholars (see e.g. Giegerich 1985; Kaye 1990; McCarthy & Prince 1990; Charette 1991; Piggott 1991, 1999; Rice 1992; Harris 1994).

The marked status of codas is observed in child language as well. Early grammars show a strong preference for CV syllables, independent of the constraints of the target language (see e.g. Jakobson 1941/1968; Ingram 1978; Fikkert 1994). When word-final consonants ultimately emerge, the predominant view is that they are syllabified as codas (e.g. Fikkert 1994; Demuth & Fee 1995; Stemberger 1996). In this paper, we argue against this view. We propose that right-edge consonants are first syllabified as onsets – specifically, as onset-nuclear sequences – regardless of their status in the language being acquired.

The principal evidence in support of this view comes from the phonetic properties that characterize early obstruent-final CVC forms: for the final consonant, the presence of aspiration (final release[1]), length, and homorganic nasal release; and for the medial vowel, the presence of length and post-vocalic pause. The data on which we base our analysis come from five English-speaking children between the ages of 18 and 26 months, shortly after the point when CVC forms emerge.

We begin in §2 by laying out our assumptions about language acquisition and about the representations that we assume for segment structure and syllable structure. We turn in §3 to demonstrate how our views on syllable structure yield a three-way typology for the syllabification of post-vocalic consonants across languages. This section concludes with a discussion of phonetic correlates of right-edge onsets in adult languages. In §4, we turn to the child data. We demonstrate that early CVC forms are characterized by release and timing properties consistent with their being syllabified as onsets. We suggest that a ban on codas is partly responsible for the shapes of these outputs: syllabifying right-edge consonants as onsets enables children to avoid building the marked representation required for a coda without compromising the segmental content of the adult target. In §5, we detail the representations that we posit for right-edge onsets. We propose that, in child language, these consonants are syllabified as onset–nuclear sequences, rather than as onsets of empty-headed syllables. In §§5.3–5.4, we draw a distinction between children who represent voicing contrasts through the Laryngeal node, as in the target English grammar, and those who initially represent this contrast through the Sonorant Voice (SV) node. We discuss how the release properties of final consonants differ as a consequence of the way that voicing is represented in the two types of grammars. Finally, in §6, we return to the syllabification of right-edge onsets in adult languages, in light of the representations that we have proposed for child language.

2. Assumptions

2.1 Acquisition

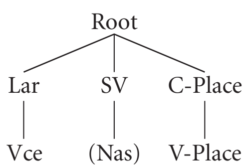
In the following lines, we outline our assumptions about acquisition. First, we adopt the view that children's early grammars are structurally unmarked (e.g. Jakobson 1941/1968; Stampe 1969; Gnanadesikan 1995). Second, we accept that children learn through exposure to positive evidence only (e.g. Chomsky

1981; Pinker 1984). In the present context, since we will argue that children initially syllabify final consonants as onsets, we must ensure that this reflects the unmarked case (see Borowsky 1986: 197–199; Piggott 1991; Goad & Brannen 2000; cf. Itô 1986: 102). In earlier work (Goad & Brannen 2000), we provided two types of arguments in favour of this position. One, following Piggott (1991), we argued that the potential for the availability of positive evidence necessitates that children begin with the analysis that final consonants are syllabified as onsets; in some languages, if learners were to start with the assumption that final consonants are instead codas, positive evidence for onset status would not be available. Two, from the point of view of processing, we demonstrated that final onsets are the optimal way to signal the right word edge, thereby yielding word-final onsets over word-final codas as unmarked. Codas are good cues to the right edge of syllables, since they depend on following onsets to license material that they cannot themselves license; however, for this same reason, they are poor cues to the right edge of words.

2.2 Segment structure

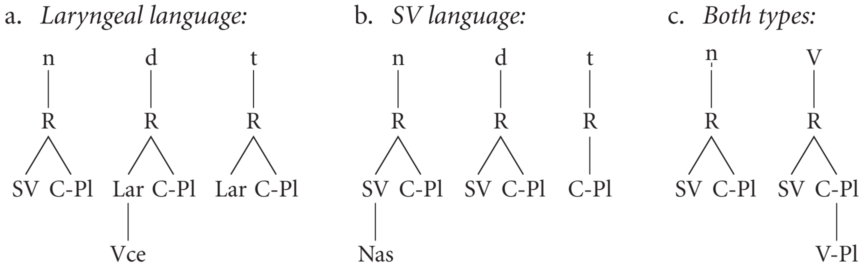
Turning to our assumptions about segment structure, we adopt the position that all phonological features are monovalent. We assume further that features are hierarchically organized as in (1) (only relevant structure is provided).

(1)



The primitives that we have selected are essential for the analysis. First, we assume that place is organized differently in consonants and vowels (Clements & Hume 1995): consonants bear a C-Place node only while vowels contain an additional dependent V-Place. Second, we assume that there is an SV node which organizes sonorant features (SV abbreviates Sonorant Voice or Spontaneous Voice; see Rice & Avery 1989; Piggott 1992). The only dependent of SV that will feature in our analysis is Nasal. It is in parentheses in (1) as we assume that in many languages this feature is not projected (Rice 1992). Indeed, in so-called Laryngeal languages, bare SV in non-nuclear position will be interpreted as nasality (Rice 1993; Avery 1996). This can be seen from the first structure in (2a). ([n], [d], and [t] abbreviate the class of nasals, voiced obstruents, and voiceless obstruents respectively.)

(2)



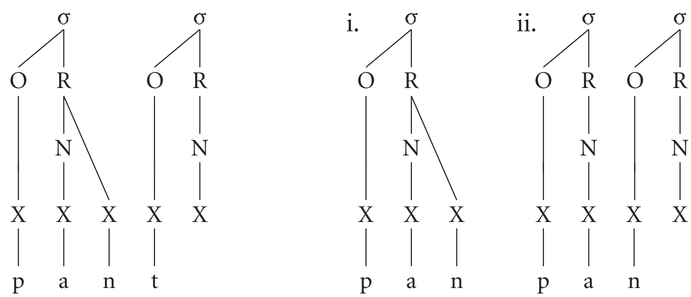
In contrast to the system in (2a), there are many languages for which there is no evidence that the Laryngeal node has been projected. These languages fall into two types. There are those for which voicing is entirely predictable and is often positionally determined (e.g. most Australian languages (Blake & Dixon 1979: 19)). We are interested in the second type: languages where voicing is formally captured through the presence or absence of the SV node, as in (2b) (see Rice 1993; Avery 1996). In the latter type of language, a bare SV node in non-nuclear position will be interpreted as a 'sonorant obstruent', e.g. [d] in (2b) (Piggott 1992; Rice 1993). Accordingly, Nasal must be projected for the class of nasals, in contrast to the system in (2a). SV languages are characterized by, for instance, the presence of allophonic variation among voiced obstruents and pre-/post-nasalized stops, nasals, and/or approximants; processes such as voicing assimilation that treat voiced obstruents and sonorants as a class; and the absence of final devoicing (Avery 1996). Turning finally to (2c), we suggest that in both Laryngeal and SV languages, a bare SV node accompanied by a C-Place node in nuclear position is interpreted as a syllabic nasal. If C-Place contains a dependent V-Place node, however, the interpretation is as a vowel. In §5, the variable interpretation of SV will feature prominently in our analysis of the release patterns that characterize early CVC forms.

2.3 Syllable structure

In this final subsection, we detail our assumptions about syllable structure. We adopt Onset–Rhyme theory with the proviso that the only sub-rhymal constituent is the nucleus. Following Government Phonology, there is no formal coda constituent (see esp. Kaye 1990; Kaye, Lowenstamm, & Vergnaud 1990: 201–202); coda consonants are organized as post-nuclear segments, as can be seen for [n] in (3a) (see also Blevins 1995: 216).[2] The principal reason for this concerns rhyme binarity. In Government Phonology, it is proposed that all sub-syllabic constituents are universally maximally binary (e.g. Kaye, Lowenstamm, & Vergnaud 1990: 199).[3] We accept a weaker version of this proposal, that constituent binarity reflects the unmarked state of affairs. If the coda were a formal constituent of the syllable, there would be no principled reason to rule out binary branching codas and, thus, ternary or even quaternary branching rhymes in the usual instance; both would be consistent with the view that sub-syllabic constituents are maximally binary.

While ternary branching rhymes are attested in languages such as Finnish, the majority of languages do not tolerate CVXC syllables word-internally. However, in several languages of the latter type, strings of this shape are tolerated at the right edge, precisely because the final consonant can be syllabified as the onset of an empty-headed syllable, as illustrated in (3a). As expected, final CC clusters in such languages have a coda–onset profile (see e.g. Rice 1992; Harris 1994, 1997; Piggott 1999; and §3.1 on Diola-Fogny).

(3) a. *Right-edge CVCC strings:* b. *Right-edge CVC strings:*



In languages which respect constituent binarity, this constraint will force the final consonant out of the rhyme in a CVXC# string. However, this is clearly not the case with CVC# strings. Following Piggott (1991, 1999), we contend that languages have two options for the syllabification of the final segment in strings of this shape: rhyme-internal consonant (coda) as in (3b.i), and onset of empty-headed syllable, (3b.ii). (A version of (3b.ii) where final consonants are syllabified as onset–nuclear sequences will be discussed in §§5–6.) We thus follow the spirit of Itô (1986) where it is proposed that languages have two choices for the syllabification of final consonants (cf. Government Phonology where final consonants are syllabified as onsets of empty-headed syllables in all languages).

Broadly speaking, there are three kinds of evidence which support two options for the syllabification of final consonants at the right edge. The first two, segmental profile and rhyme shape, appear in the literature; the third, release properties, will be explored in this paper, mostly within the context of child language.

i. Segmental profile: In languages where final consonants are syllabified as codas, these segments have a coda profile in the sense that the restrictions observed mirror those that hold of word-internal codas. Languages which fall into this category include Selayarese and, under most analyses, Japanese (e.g. Itô 1986; cf. Yoshida 1990). In languages where final consonants are syllabified as onsets, on the other hand, these segments have an onset profile in the sense that there are (virtually) no restrictions on what can be present in this position as compared to what can appear in word-internal coda position in the same language. Diola-Fogny and French fall into this class of languages.
ii. Rhyme shape: In languages where final consonants are syllabified as onsets and branching rhymes are permitted, strings of the shape CVXC are tolerated in final position as in, for example, Diola-Fogny (CVVC and CVCC) and Yapese (CVVC). This is in direct contrast to languages where right-edge consonants are syllabified as codas; in such languages, final strings are limited to CVC (and CVV, if licit), as in Selayarese. In some of these languages, claims about rhyme shape can be supported through alternations: closed syllable shortening, vowel epenthesis, consonant deletion, etc.
iii. Release properties: Right-edge consonants which are syllabified as onsets may be, but are not required to be, characterized by release properties similar to those displayed by onsets that are followed by phonetically realized vowels. Yapese is a representative example. In this language, final plain stops are 'aspirated' (Jensen 1977: 27).[4] We will see in §4 that CVC forms in early child language resemble Yapese in this respect. By contrast, final codas do not display such release properties; indeed, they are often subject to neutralization constraints which prohibit the appearance of such properties.

3. Right-edge consonants in adult languages

3.1 Typology for the syllabification of post-nuclear consonants

In this section, we will briefly demonstrate how the assumptions outlined above serve to limit the options available for the syllabification of post-nuclear consonants. A three-way typology will result. As mentioned in §2.3, the factors which are typically used to determine the status of right edge consonants are segmental profile and preceding rhyme shape. We will look at three languages from this perspective: Selayarese, where word-final consonants are syllabified as codas, and Diola-Fogny and Yapese, where final consonants are syllabified as onsets. Diola-Fogny differs from Yapese in that, like Selayarese, Diola-Fogny permits word-internal codas. Our discussion of these three languages draws heavily on Piggott (1999).

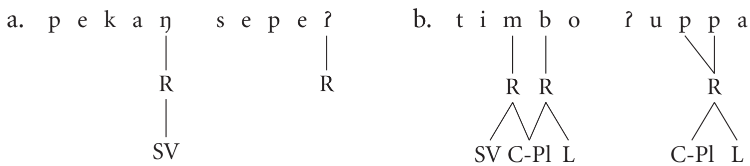
We begin with Selayarese where word-internal codas are restricted to geminates, homorganic nasals, and [ʔ]; see (4a) (transcriptions simplified from Mithun & Basri). Word-finally, consonants have a coda profile: only [ŋ] and [ʔ] are found, (4b). Further evidence for this comes from Mithun and Basri's observation that words which end in consonants other than [ŋ] and [ʔ] are subject to final epenthesis (e.g. /katal/ → [katala] 'itch', /tulis/ → [tulisi] 'write').

(4) *Selayarese (Austronesian)* (Mithun & Basri 1986):

a. ʔuppa 'find' ʔandeŋka 'throw' laʔba 'lack of salt'
allonni 'this day' timbo 'grow' seʔla 'salt'

b. pekaŋ 'hook' sepeʔ 'narrow passage'
potoŋ 'style' sassaʔ 'lizard'

If [ŋ] and [ʔ] are placeless – [ŋ] because the least marked nasal is interpreted as velar in many languages (Trigo 1988; Rice 1996), and [ʔ] because it lacks supralaryngeal constriction – we arrive at the representations in (5a) where the Selayarese coda can only license nasality, a bare SV node in our approach (see Piggott 1999; cf. Goldsmith 1990). Word-internal codas also abide by this restriction, but as they can share features, which they cannot themselves license, with a following onset, a wider range of options is observed in this context, (5b).

(5)


On the face of it, Diola-Fogny is similar to Selayarese: word-internal codas are limited to nasal geminates or to homorganic nasals and liquids, (6a). The second column in (6a) reveals that consonant deletion due to morpheme concatenation yields forms which are consistent with this. In contrast to Selayarese, though, any consonant from the Diola inventory can occur in final position (Sapir 1965: 6). Representative examples are provided in (6b).

(6) *Diola-Fogny (West Atlantic)* (Sapir 1965):

a.	niŋaŋŋan	'I cried'	takʊn – mbɪ	→ takʊmbɪ	'must not . . . '	
	jɛnsʊ	'undershirt'	lɛt – kʊ – jaw	→ lɛkʊjaw	'they won't go'	
	saltɛ	'be dirty'	na – laŋ – laŋ	→ nalalaŋ	'he returned'	

b.	fumoːməf	'the trunk'	ufeːgir	'three'	famb	'annoy'
	wopuːs	'green caterpillar'	irok	'I am fat'	jawac	'to swim'

According to Itô (1986), word-final consonants in Diola-Fogny are not codas, but are instead licensed by extraprosodicity. More recently, Piggott (1991, 1999) has demonstrated the parallel between the segmental profile of word-final consonants and what are indisputably onsets in this language; he thereby concludes that final consonants are onsets of empty-headed syllables. We concur with this analysis. Further support comes from the coda–onset profile of final clusters (e.g. [famb]) and from the absence of closed syllable shortening (e.g. [wopuːs]).

In sum, Selayarese and Diola-Fogny are consistent with the hypothesis that codas are marked vis-à-vis onsets; codas typically license a limited range of features with the consequence that there are severe restrictions on what can appear in this position. By contrast, onsets are (virtually) unrestricted across languages. The range of options permitted among right-edge consonants in Diola is consistent with their falling into the latter category.

The unmarked status of onsets is further demonstrated by the fact that no language prohibits CV syllables and some languages only allow for syllables of this type (Cayuvava, Senufo). Yapese, the last adult language we will consider, is similar to Cayuvava and Senufo in that it forbids word-internal codas. Like Diola-Fogny, however, word-final consonants (barring [h]) can freely occur (Jensen 1977: 29). (In fact, words must end in a consonant; we will return to this in §6.) Other languages with the same restrictions on syllable shape as Yapese include Luo (Stafford 1967) and Kamaiurá (Everett & Seki 1985).

Representative examples of Yapese word-final onsets are given in (7a). The illicit forms in (7b) reveal the absence of word-internal codas. Indeed, epenthesis serves to break up clusters which arise through morpheme concatenation: /bug-y/ → [buguy] 'to bend something', /meel-y/ → [meeliy] 'to pull on rope'.[5]

(7) *Yapese (Austronesian)* (Jensen 1977):

a.	laaṭ	'type of tree'	magad	'lime container'	taaŋ	'song'
	garik	'stinging jellyfish'	pilig	'to take down'	lik'	'its root'
	lukur	'stick to pick up food'	danoop	'the world'	faraf	'floor'

b. *lukkur *piltig *dandoop *fardaf

The lack of word-internal codas, the presence of long vowels before word-final consonants, and the range displayed in the segmental shape of final consonants all strongly suggest that these consonants are not syllabified as codas. Piggott (1991, 1999) proposes that they are onsets of empty-headed syllables. Thorburn (1993), by contrast, suggests that they are extraprosodic. We concur with Piggott that final consonants in Yapese are onsets; however, there are some differences between our analysis and his which will be addressed in §6.

In sum, we have arrived at a three-way typology as concerns the syllabification of post-nuclear consonants; this typology is provided in Table 1. Languages with word-internal codas have two options for final consonants: they may be syllabified as onsets (Diola-Fogny), or they may be codas (Selayarese). Since the shape of word-internal codas plays an important role in determining the status of word-final consonants, we predict that in languages with final consonants but no word-internal codas, final consonants will always be analysed as onsets (Yapese). In other words, there should be no language where CVC is limited to the right word edge and where final consonants have a coda profile or display coda-like behaviour. Under our approach, if final consonants were syllabified as codas in such a language, there would be nothing to prevent the presence of word-internal codas.

Table 1. Typology for the syllabification of post-nuclear consonants

Word-internal codas	*Word-final consonants*	*Example languages*
Yes	Onset	Diola-Fogny, French
Yes	Coda	Selayarese, Japanese
No	Onset	Yapese, Kamaiurá
No	Coda	–

English, the language being acquired by the children under present investigation, has been noticeably absent from the discussion thus far. In some respects, it falls into the same category as Selayarese and Japanese: final consonants in CVC strings are syllabified as codas. Evidence for this comes from word-minimality effects: the minimal monosyllabic word is CVX, indicating that the final consonant in CVC is weight-bearing. However, in contrast to Selayarese and Japanese, English tolerates CVXC strings at the right edge (we ignore the extra position reserved for [t,d,s,z] inflection). In words of this shape, final consonants must be syllabified as onsets of empty-headed syllables given that, with limited exceptions (see Harris 1994), word-internal rhymes in English abide by constituent binarity discussed in §2.3. Closed syllable shortening supports this analysis; compare [faɪ.v] 'five' with [fɪf.tiː], *[faɪf.tiː] 'fifty'.[6]

3.2 Phonetic correlates of final onset status

Thus far, the evidence we have used to determine the status of final consonants has been distributional. The question which we turn to now is whether there are any phonetic correlates of onset versus coda status of these consonants. We propose that word-final consonants which are syllabified as onsets may be characterized by release properties similar to those observed for onsets which are followed by phonetically realized nuclei. Yapese exemplifies this situation: plain voiceless stops in final position are 'aspirated' (Jensen 1977: 27). Yapese is not anomalous in this respect: Sierra Popoluca (Elson 1947) and Nez Perce (Hoard 1978) also display aspiration of onsets followed by phonetically-empty nuclei.

Importantly, we are not claiming that, in the absence of an appropriate phonetic correlate, final consonants are syllabified as codas. The stronger claim cannot be made, as there may be overriding constraints in a language. Specifically, languages may limit the licensing options of consonants at the right word edge, regardless of how such consonants are syllabified.[7] For example, we have seen that final consonants in Diola-Fogny are syllabified as onsets; yet Sapir (1965: 5) notes that voiceless stops in this position are optionally unreleased. We return to this issue in §6, after we have provided the representations that we posit for onsets at the right word edge.

As concerns final codas, we suggest that it would be highly unlikely for a release property such as aspiration to be systematically present on a coda. Indeed, neutralization is typically observed in this position, and consonants which undergo laryngeal neutralization are often unreleased (for various approaches to laryngeal neutralization, see e.g. Westbury & Keating 1986; Cho 1990; Lombardi 1991; Harris 1997; Steriade 1997; Kawasaki 1998).

4. CVC forms in early child language

We turn now to the phonetics of CVC forms in early child language. We will see that child language resembles Yapese, as final consonants are characterized by release properties like aspiration. The parallel exists on distributional grounds as well: in child language, word-final consonants emerge before there are any word-internal codas. Thus, at the stage that we are focussing on, the only postvocalic consonants are in word-final position. We propose that final consonants in early CVC forms are not syllabified as codas; indeed, they cannot be under the claims made in Table 1 above. We will argue that the release and timing properties of early obstruent-final CVC forms are consistent with their being syllabified as ON.ON where the features of the right-edge consonant are shared across the final onset–nuclear sequence.

4.1 Epenthesis, truncation, and word minimality

We begin with epenthesis and truncation, two processes that are standardly used as evidence for an early preference for open syllables. Representative data are given in (8).[8] In the case of epenthesis, the epenthetic vowel may surface as a copy of the preceding vowel (8a.i), or as some default vowel (8a.ii). In both cases, as with truncation in (8b), the result is a string of open syllables. Both epenthesis and truncation are thus consistent with a constraint such as NoCoda (Prince & Smolensky 1993), technically NoBranchingRhyme in our approach, being undominated in children's early grammars.

(8)	a. *Epenthesis:*				b. *Truncation:*			
	i.	[ʌbʌ]	'up'	(Jacob 20 mos)		[bæ]	'bath'	(Hildegard 20 mos)
		[hɛtɛ]	'hat'	(Jacob 20 mos)		[be]	'Bates'	(Hildegard 20 mos)
	ii.	[bɪbi]	'bib'	(Mollie 18 mos)		[ko]	'coat'	(Mollie 18 mos)
		[wåki]	'walk'	(Mollie 18 mos)		[dɔ]	'dog'	(Mollie 18 mos)

Abiding by NoCoda through truncation, however, yields forms which are subminimal. This is clearly true for words such as [bæ] where the vowel is short in the adult grammar. In fact, we contend that all the forms in (8b) are subminimal. First, vowel length is variable in the early outputs of children learning West Germanic languages, suggesting that it is not initially used contrastively (Fikkert 1994; Demuth & Fee 1995). Second, given that the quality difference observed among English vowel pairs – [iː-ɪ], [eː-ɛ], etc. – is very salient, it seems reasonable to propose that children initially analyse the difference in long–short pairs at the melodic level only, with a feature such as [tense]. If all truncated forms at this stage are then subminimal, the question which must be addressed is whether the child 'knows' this. We think not, as there is no evidence that the child has any knowledge of the Foot at this stage in development (Demuth & Fee 1995; Goad 1997). Consider for example Mollie. According to Holmes (1927: 221), *all* of Mollie's two-syllable outputs at the stage we are focussing on received equal stress on both syllables. While such forms could be parsed as two degenerate feet, $(CV)_{Ft}(CV)_{Ft}$, the unmarked role that Foot-Binarity plays across languages makes this highly unlikely, especially in light of the commonly held view that children's early grammars reflect unmarked properties (cf. §2.1). If, by contrast, the Foot has not yet been projected, the monotonic stress pattern described by Holmes (and others) is exactly as expected. Following from this, we propose that FtBin plays no role in the emergence of right-edge consonants under our analysis; accordingly, there is no constraint to initially drive children to syllabify final consonants as codas and they are free to syllabify them as onsets (cf. Fikkert 1994; Demuth & Fee 1995).

4.2 The phonetics of final consonants

As mentioned earlier, when final consonants appear in children's outputs, they display a number of effects which suggest that they are syllabified as onsets, regardless of the constraints of the target grammar. We begin in this section by discussing the release properties of final obstruents in early CVC forms. In §4.3, we turn to properties of the medial vowel in CVC forms.

The examples in (9) show that final stops may surface as aspirated (finally released). Fey and Gandour (1982: 74) describe Lasan's right-edge stops as having a "distinctive oral release". Leopold (1939: 108) mentions that, at 22 months of age, Hildegard's right-edge stop in 'meat' was "strongly aspirated with the exaggeration typical of the first final consonants".

(9) *Final aspiration:*

[dap^h]	'drop'	(Lasan 21–25 mos)	[vit^h]	'feet'	(Lasan 21–25 mos)
[mit^h]	'meat'	(Hildegard 22 mos)	[bok^h]	'broke'	(Hildegard 22 mos)
[bak^h]	'bike'	(Jacob 20 mos)	[ap^h]	'up'	(Jacob 20 mos)

If these consonants are syllabified as codas, the data seem somewhat anomalous when compared with adult languages. As mentioned earlier, in adult languages, laryngeal neutralization is commonly observed in coda, and neutralized stops are often unreleased. Westbury and Keating (1986) and Steriade (1997) provide phonetically-based accounts of laryngeal neutralization in final position, the former based on ease of articulation, and the latter, on the observation that fewer cues to voicing contrasts are available in this position. If child language reflects the unmarked state of affairs, and markedness is partly determined by articulatory and/or perceptual factors, one might wonder why we observe final aspiration in children's early outputs. Under our analysis, an explanation readily emerges: final consonants are syllabified as onsets, more specifically, as onset–nuclear sequences.

In (10), we can observe the second pattern for final consonants in CVC forms: the consonants may exhibit lengthening. With regard to Mollie's /k/-final forms, Holmes states that "the explosion of *k* was ... prolonged (almost equal to a *schwa*)" (p. 221). Thus, the release, rather than the closure, was twice as long, revealing a striking parallel between the pattern in (10) and what was described as final aspiration in (9). In §5.2, we argue that final aspiration and final length are in fact equivalent.

(10) *Length on final consonant:*

[keˑkː]	‘cake’	(Mollie 18 mos)	[dʒusː]	‘juice’	(Jacob 20 mos)
[æˑtː]	‘hat’	(Mollie 18 mos)	[s̀iːs̀ç]	‘cheese’	(Jacob 20 mos)
[bæˑdː]	‘bad’	(Mollie 18 mos)			

The data in (11) show that right-edge voiced stops may be followed by a homorganic nasal release. Fey and Gandour (1982: 74) report that, from 21–25 months of age, Lasan’s “word-final voiced stops ... were consistently produced ... with a distinctive nasal release”. In §5.3, we will demonstrate that this CN̩ pattern is not anomalous; on the contrary, it is exactly what is expected for children who represent voicing contrasts with the feature SV rather than Laryngeal. For the present, notice that Fey and Gandour transcribe the nasal release as syllabic; they provide evidence for this analysis from the pitch contour (p. 72). Importantly, if the nasal is syllabic, then the root-final consonant is optimally syllabified as an onset, not as a coda.[9]

(11) *Nasal release:*

[dabm̩]	‘stub’	(Lasan 21–25 mos)	[babm̩]	‘bulb’	(Lasan 21–25 mos)
[vidn̩]	‘feed’	(Lasan 21–25 mos)	[dædn̩]	‘dad’	(Lasan 21–25 mos)
[bəgŋ̍]	‘bug’	(Lasan 21–25 mos)	[wɔgŋ̍]	‘frog’	(Lasan 21–25 mos)

4.3 The phonetics of pre-consonantal vowels

In this section, we consider two properties that characterize the medial vowel in CVC forms. These properties are consistent with the vowel being at the right-edge of a syllable, an analysis which would only hold if the final consonant were syllabified as an onset. We begin with length on the medial vowel in (12), either half length as seen in Mollie’s examples, or full length as in Jacob’s examples.

(12) *Vowel length:*

[keˑkː]	‘cake’	(Mollie 18 mos)	[ʌmʌːk]	‘milk’	(Jacob 20 mos)
[æˑtː]	‘hat’	(Mollie 18 mos)	[s̀iːs̀ç]	‘cheese’	(Jacob 20 mos)
[bæˑdː]	‘bad’	(Mollie 18 mos)			

Vowel length is not unexpected if the vowel is at the right edge of a syllable, but it would be highly unexpected if the following consonant closed the syllable. This is especially true in examples like [keˑkː] and [ʌmʌːk], as vowels tend to be shorter before tautosyllabic voiceless consonants in English (e.g. Ladefoged 1993:90). Notice that many of the forms in (12) are repeated from (10); it seems that children often display length on both the medial vowel and on the final consonant. As will be seen in §5, this follows from the syllabification that we propose.

The data in (13) exhibit the final pattern to be accounted for: VC sequences may be interrupted by a pause (marked by a period in Scott's data and by a hyphen in Hildegard's and Jacob's data[10]). Post-vocalic pause is expected if the vowel is at the right edge of a syllable, and the final consonant is the onset of the following syllable. This will be discussed further in §5.1.

(13) *Post-vocalic pause:*

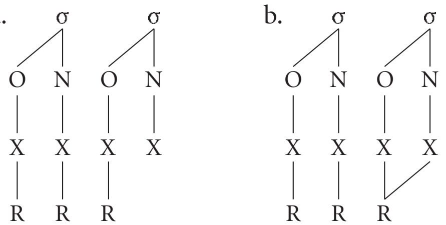
[sɚ.s]	'stairs'	(Scott 23 mos)	[be-t^{h}]	'Bates'	(Hildegard 19 mos)
[ɕu.s]	'shoes'	(Scott 23 mos)	[ba-kh]	'box'	(Jacob 19 mos)
[s^{j}u.s]	'juice'	(Scott 23 mos)			

We began §4 by providing data from epenthesis and truncation for target CVC, and we suggested that undominated NoCoda was partly responsible for these patterns. We propose that this represents the first stage in development. At Stage 2, our focus in this paper, CVC target forms are realized as such. However, we have seen that they display a variety of timing and release properties which are not manifested in the target grammar. We have argued that all of these patterns can be accounted for if the final consonant is not syllabified as a coda. As such, these forms constitute further support for the dominant role played by NoCoda in early acquisition.[11] Syllabifying final consonants as onsets (onset–nuclear sequences) enables children to avoid the marked representation required for a coda consonant, a branching rhyme, without severely compromising the segmental content of the adult target. This is particularly apparent for a child who at Stage 1 displays truncation, CVC → CV. At Stage 2, target CVC is realized as CV.C and the child can thereby add melodic content to his/her Stage 1 output without adding complexity to the syllable structure representation.

5. Representations

In this section, we detail the representations that we posit for early CVC forms. Since there are no codas (branching rhymes) at this point in development, syllables are limited to (O)N. Consistent with this, there are two options for the syllabification of CVC inputs: one, final consonants are onsets of empty-headed syllables (14a); or two, final consonants are syllabified as onset–nuclear sequences (14b). (In the interest of space, rhyme nodes have been eliminated; R abbreviates Root nodes and all dependent features.)

(14) a. b.



For input CVC with the output representation in (14a), there is no overt nuclear material following the final consonant. Accordingly, the release properties discussed in §4.2 would be the phonetic manifestation of the right-edge consonant being syllabified as an onset. In (14b), by contrast, the release properties would be present precisely because the melody of the final consonant has spread into the nucleus; in other words, input CVC would be realized as $(\text{C+V})_\sigma(\text{C+Release})_\sigma$. We will see shortly that only the representation in (14b) enables us to unify the phonetic properties of final consonants observed in children's outputs – aspiration, final length, and homorganic nasal release. Before we turn to this, we will demonstrate that the phonetic properties of pre-consonantal vowels – vowel length and post-vocalic pause – both follow straightforwardly from a configuration where the final consonant in a CVC target is syllabified as an onset, rather than as a coda.

5.1 Vowel length and post-vocalic pause

Recall from §4.3 that, for some children, the medial vowel in a CVC target may appear with (half) length. Indeed, all of Mollie's CVC outputs during the period under investigation display half length, e.g. [æˑtː] 'hat'. For other children, we had observed that CVC forms may appear with post-vocalic pause as in, for example, Hildegard's [be-t^h] 'Bates' and Jacob's [ba-kh] 'box'. We suggest that both of these effects follow from the structures in (14) where the final consonant is syllabified as an onset. Concerning vowel length, the crucial consequence of such a representation is that the initial syllable is open. As vowels tend to be longer in open syllables across languages, length in this position in the children's outputs is entirely as expected. Concerning medial pause, the relevant property of the representations in (14) is that the vowel and final consonant are not internal to the same constituent. If the VC string in a CVC target were syllabified as a rhyme, post-vocalic pause would be highly unexpected as it would interrupt a sub-syllabic constituent. By contrast, under the syllabifications in (14) where the final consonant is an onset, a major constituent boundary coincides with the pause.

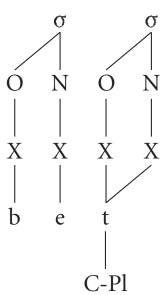
As we have seen, vowel length and post-vocalic pause are consistent with both of the structures provided in (14): the final consonant is syllabified as an onset and is thereby in a constituent separate from the preceding vowel. These patterns do not, however, speak to whether the final consonant should be syllabified as the onset of an empty-headed syllable (14a), or as an onset–nuclear sequence (14b). As mentioned above, when the phonetic properties of final consonants are taken into consideration, only (14b) allows for a unified account of the various patterns observed. This will be demonstrated in the following subsections. Finally, as the structure in (14b) has consequences for the representation of right-edge onsets in adult languages, we will turn to this issue in §6.

5.2 Aspiration and final length: Voiceless outputs

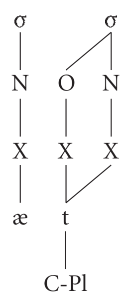
The structure in (14b) reflects the fact that what is often described as final aspiration is more accurately called final *release* (cf. Note 1). Consistent with this, we provide the representation for final aspiration in (15a). (Segments abbreviate Root nodes; only relevant features are provided throughout.)

(15)

a. *Final aspiration: Hildegard's [be-t^h]:*



b. *Consonant length: Mollie's [æˑtː]:*



Recall from the data in (10) that Mollie's voiceless stops are transcribed by Holmes (1927) as long. While geminates are disfavoured at edges cross-linguistically, length as final release is exactly as expected if these 'geminates' are syllabified as onset–nuclear strings, as in (15b): length is audible precisely because the consonant is released as a nucleus. A comparison of (15a) and (15b) reveals that, for voiceless stops, we have provided the same representations for aspiration and final length. We believe that they are transcriptional variants of the same phenomenon. (We have not found any author who transcribes some final voiceless stops with aspiration and others with length.) Both may be described as a homorganic burst of noise.

If the input-final consonant is partly syllabified in the nucleus, one might wonder why the release does not sound vowel-like, in particular, like a voiceless schwa. We contend that this is because final aspiration is not represented in the same manner as is voiceless schwa (cf. e.g. Urbanczyk 1996; see also Dyck 1990). To elaborate on this, we must return to (2c) where we proposed that vowels contain SV and V-Place. Given that SV refers to *spontaneous* voicing, we propose that voiceless vowels lack SV; however, they retain the V-Place specification of the corresponding voiced vowel. Voiceless schwa, being placeless, is represented as a bare V-Place node. The structure in (15a) reveals that final aspirated stops, in contrast to voiceless vowels, do not have a V-Place node; thus, although they are syllabified in the nucleus, they are still contoids.

Our representation of the final aspirated stop as a 'geminate' ON sequence means that the final release should sometimes be audible as length on the consonant, as in Mollie's stops and Jacob's fricatives in (10), or as a fricative homorganic with the preceding stop, as in some of Hildegard's forms: " … the aspiration [in 'meat' at 22 mos] sometimes even took the form of a homorganic fricative, [ç]" (Leopold 1939: 108). Similar observations have been made for adult languages. Regarding right-edge aspirated ('syllabic') stops in Bella Coola, Hoard (1978: 72) states: "The release or 'burst' that accompanies a syllabic stop … does not resemble very closely a voiceless vowel … . One does of course hear the frequencies characteristic of each of the consonantal bursts, but these are not the frequencies associated with the vowel formants of voiceless vowels." Our geminate-like structure for right-edge aspirated stops is very similar to that proposed by Hoard (1978).[12]

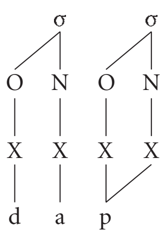
In the absence of spectrographic analysis to substantiate our claims about the representation of children's final aspirated consonants, we must rely on other evidence. The epenthesis patterns seen in (8a) offer some support for our view: vowels inserted to satisfy NoCoda are always voiced, even when the epenthesized vowel is schwa (Hildegard at 26 mos: [dɛtə] 'get', [dotə] 'don't'). If final aspiration is interpreted as schwa-insertion, one must wonder why schwa is sometimes voiceless (i.e., aspiration) and at other times voiced. The quality of adjacent segments cannot be responsible for this, as the following forms produced by Hildegard at 26 months reveal: [ʔɛtʰmaɪhu] 'in my room' versus [gɔtəmaɪʒu] 'got my shoe'. We thus reiterate our position: aspiration is not represented in the same fashion as voiceless schwa.

5.3 SV voicing versus laryngeal voicing

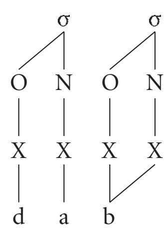
We turn now to consider Lasan's grammar in some detail. Lasan is one of the children who always produces final aspiration for input voiceless stops. On the face of it, voiced stops show a different pattern. Recall that they are produced with nasal release. Lasan's outputs for 'drop' [dap^h] and 'stub' [dabm̩] from (9) and (11) are repeated in (16), along with the representations that we propose.

(16)

a. *Final aspiration: Lasan's [dap^h]:*



b. *Nasal release: Lasan's [dabm̩]:*



(16b) reveals that we do not consider Lasan's nasally-released outputs to be anomalous; they are represented in the same fashion as are final aspirates and long consonants. Two facts mentioned earlier suggest a structure along these lines: 1. the nasal release is syllabic and so must occupy the nucleus; 2. the nasal is always homorganic with the preceding stop, indicating that the string minimally shares place features. To answer the question as to why these forms surface with a *nasal* release, however, we must look at other properties of Lasan's grammar.

Voicing contrasts for Lasan's obstruents are neutralized word-initially: [bɪgŋ̩] 'pig' *and* 'big'. Stops display a lot of variation in this position, in contrast to what is observed finally. The variation reveals both voiced obstruent- and sonorant-like outputs: [^{n}du ∼ ɗu ∼ du ∼ d̥u] 'two', [bədn̩ ∼ mədn̩ ∼ βədn̩] 'bird' (Fey & Gandour 1982: 73). Recall that similar alternations among sonorants and voiced obstruents are typical of SV languages (cf. §2.2). Laryngeal is not projected in these languages, and voicing contrasts are captured through SV. We propose that the patterns of variation exhibited by Lasan's initial stops, as well as the realization of his final voiced stops as CN̩ sequences, reveal that his grammar also expresses voicing contrasts with SV. Instead of Laryngeal voicing which is used by the target English grammar, Lasan makes a contrast between (voiced) sonorant obstruents which bear an SV node, and (voiceless) ‘obstruents’ which are not specified either for Laryngeal or for SV. See (17).[13,14]

(17)

a. *Lasan’s grammar: SV language:*

b. *Target English: Laryngeal language:*

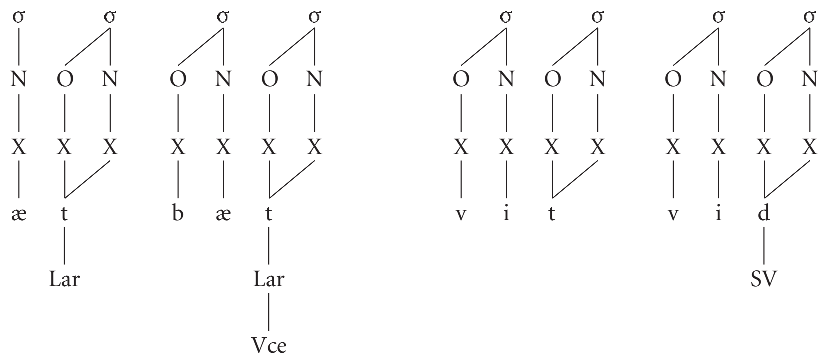
One might question why Lasan would analyse English as an SV language. Two facts suggest that SV has a privileged status across languages. One, all languages contrast sonorants and obstruents, but not all have laryngeal contrasts; in the latter case, languages may capture voicing through SV, or voicing may be entirely predictable with no sonority node required on consonants (cf. §2.2). Two, the sonorant–obstruent contrast is acquired before laryngeal contrasts are acquired (e.g. Jakobson 1941/1968; Shvachkin 1948/1973; Smith 1973). On the basis of these observations, we hypothesize that SV is available to the learner before Laryngeal is available. Initially, SV is required to capture the contrast between sonorants (which bear this node) and obstruents (which bear no sonority node). As the child’s set of contrasts increases, Laryngeal will be projected if required by the target language, but if the need for Laryngeal has not yet been detected, voicing contrasts can be captured through SV. This reflects Lasan’s grammar.[15]

Lasan can be compared with other children in this study who express voicing contrasts with Laryngeal. Mollie, for instance, has a voicing contrast initially (e.g. [tʌˑkː] ‘Tuck’ vs. [dʌˑkː] ‘duck’) as well as finally (e.g. [æˑtː] ‘hat’ vs. [bæˑdː] ‘bad’), and no sonorant-like outputs are observed in either position. Final stops surface as long, consistent with their being syllabified as ON sequences in her grammar, as they are in Lasan’s grammar. The different ways that final stops are realized in these two types of systems is illustrated in (18) (relevant structure only). For voiceless stops, length/aspiration will occur in both types of grammars because of the ON syllabification of final consonants. This will be true whether these consonants are specified for Laryngeal (18a), or unspecified for a sonority node (18b). A difference in surface forms is evidenced for input voiced stops only. In a Laryngeal language, these will be realized as long whereas in an SV language, they will optimally surface with a nasal release. While in the latter case, they could also surface as long, [vidː], we propose that the favoured interpretation of an SV consonant in the *nucleus* is a nasal, not a stop (cf. (2c)). One might question further why we do not find a vowel in place of the nasal; we address this issue next.

(18)

a. *Laryngeal language (Mollie):* [æˑtː] 'hat' [bæˑtː] 'bad'

b. *SV language (Lasan):* [vit^{h}] 'feet' [vidn̩] 'feed'



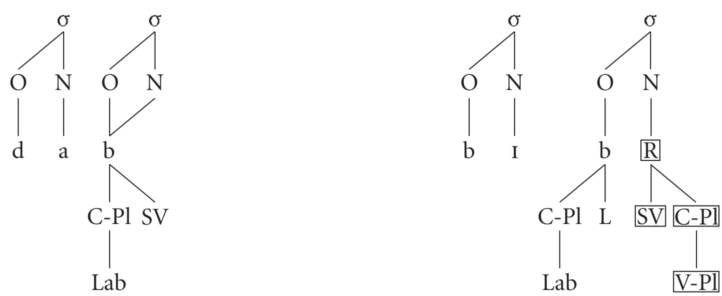
5.4 Variable interpretation of SV

We have argued that homorganic CN̩ clusters arise from final voiced stops which are syllabified as onset–nuclear sequences in an SV language. Given that SV is a feature that both consonants and vowels can bear, one might ask why the presence of SV in nuclear position does not yield a vowel in the case of (18b). We do assume that an SV node without dependents is interpreted as a vowel in nuclear position (cf. (2c)), as will be seen shortly when we discuss final epenthesis. Accordingly, if both CN̩ clusters and vowel epenthesis arise as a consequence of a ban on codas (cf. §4), we must ensure that the variable interpretation of nuclear SV can be predicted from some property of the representation.

Recall from §5.3 that Lasan's CN̩ clusters are always homorganic: [dabm̩], *[dabn̩]. Further, the data in (11) reveal that he has acquired a range of C-Place contrasts: [dabm̩], [vidn̩], [bəgŋ̍]. We can thus conclude that Lasan's CN̩ clusters share C-Place and dependent features. The structure in (19a) demonstrates that it is the combination of SV and C-Place (without V-Place) that yields the interpretation of the final nucleus as a nasal (cf. Rice 1992).

Turning to the case of epenthesis in (19b), where 'bib' → [bɪbi], one can see that, in contrast to (19a), C-Place is not shared between the epenthesized vowel and the preceding consonant (inserted material is boxed). Consequently, nuclear SV is interpreted as a vowel (cf. (2c)). We propose that vowel epenthesis is primarily driven by two constraints, Nuc ⊃ SV and Nuc ⊃ V-Place, which together reflect the fact that nuclei are in the unmarked case voiced vowels. In Mollie's grammar, where the epenthetic vowel is realized as [i], other features (not shown in (19b)) must be inserted in order for the vowel to receive its interpretation. We suggest, however, that Mollie's cases of epenthesis are remnants from an earlier stage in development. Her current grammar, like that of Lasan, shows a preference for ON feature sharing. This enables her to produce outputs that are more faithful to the input segmental content while at the same time abiding by NoCoda.

(19)
a. *Nasal release: Lasan's [dabm̩]:* b. *Vowel epenthesis: Mollie's [bɪbi]:*



6. Representation of right-edge onsets in adult languages

Before concluding, we will briefly address the status of final consonants in adult languages such as Diola-Fogny and Yapese. An issue that arises is whether our analysis of right-edge onsets in child language as onset–nuclear sequences forces us to make the same claim about the syllabification of right-edge onsets in adult languages. The alternative is that in (some) adult languages, right-edge onsets are syllabified as onsets of empty-headed syllables, as proposed in much of the literature (references in §1). At this early point in our investigation, we see no reason not to allow for both options. We do, however, address two aspects of this issue here. What kind of constraint would favour the onset–nuclear syllabification in child language? What type of evidence could we use to determine which option a particular adult language employs?

With regard to the first question, we suggest that there is a constraint which disfavours empty nuclei and that this constraint is initially undominated. Together with NoCoda, this will ensure onset–nuclear syllabification for CVC strings for children whose grammars are faithful to the segmental content of the input. For grammars that do not tolerate ON feature sharing, the result will be epenthesis (8a) or truncation (8b).

With regard to the second question, it may be the case that the presence of an onset-like release such as aspiration indicates that the final consonant has 'spread' into the nucleus (cf. Hoard 1978). Thus, in Yapese, final consonants would be syllabified as onset–nuclear sequences, while in Diola-Fogny, they would be onsets of empty-headed syllables. At first glance, an additional fact about Yapese, discussed at length in Piggott (1999), would appear to suggest that this analysis is not correct: words in this language must end in consonants; final vowels are subject to apocope. If final consonants are syllabified as ON sequences in Yapese, we must explain why vowels at the right edge appear to delete. We contend that while the melodic content of these final nuclei deletes, the nuclear position and X slot remain. Melody deletion is motivated by cues to word edge. Specifically, recall from §2.1 that we had mentioned that in Goad and Brannen (2000), we argue that the best way to signal the right edge of a word is to end the word in an onset (or onset–nuclear sequence). Since Yapese only has open syllables word-internally, the constraint that words must end in a consonant is a particularly good cue to the right edge of this domain.

7. Conclusion

In conclusion, we have proposed that children's first word-final consonants are syllabified as onset–nuclear sequences, not as codas as is standardly assumed. First, we argued that distributional evidence supports the view that NoCoda is still operative when CVC forms emerge: word-final consonants appear in development when word-internal codas are absent; drawing on parallels with adult languages, we demonstrated that the final consonant in a CVC# string would thus have to be an onset.

Second, we demonstrated that children's first CVC forms display release and timing properties – aspiration and/or length on the final consonant, homorganic nasal release, vowel length, and post-vocalic pause – which do not support the traditional claim that final consonants are syllabified as codas. Instead, we proposed that they motivate an analysis where a right-edge consonant is syllabified as the onset of an independent syllable. We argued further that aspiration/length on the final consonant and homorganic nasal release support a representation where this onset additionally shares its features with the following nucleus. We demonstrated that the advantage for the developing grammar is that syllabifying final consonants as onset–nuclear sequences allows children to avoid building the marked representation required for a coda, while at the same time permitting them to be more faithful to the segmental content of adult CVC inputs.

Finally, we demonstrated that in some adult languages where right-edge consonants are syllabified as onsets, relevant phonetic properties may be systematically present on these segments as well, for example, aspiration in Yapese. We speculated that in such languages, final onsets may be syllabified as onset–nuclear sequences, as in early child language, and thereby suggested that there may be two options for the syllabification of right-edge onsets in adult languages: onset–nuclear sequences and onsets of empty-headed syllables.

We have focussed on an early stage in production when CVC forms are first emerging. We have seen that a diverse range of consonants which are accompanied by onset-like releases are observed in final position at this stage. A logical next step in the investigation would be to look at subsequent stages in development when children begin to syllabify singleton final consonants (after short vowels) as codas in English-type languages. In these later stages, we would expect final consonants to observe restrictions characteristic of codas, and thus, to display accompanying effects such as debuccalization. Some preliminary evidence that this expectation is met can be found in Goad and Brannen (1998).

Notes

* Earlier versions of this paper were presented at the 4th HILP Conference, at the Annual Meeting of the Canadian Linguistic Association, at the Montréal-Ottawa-Toronto Phonology Workshop, at University College London, and at the University of Essex. We would like to thank the audiences for questions related to this work. We would especially like to thank Evan Mellander, Joe Pater, Keren Rice, Iggy Roca, Kevin Russell, and an anonymous reviewer for detailed comments on an earlier written version. This research was supported by grants from FCAR (to H. Goad) and SSHRC (to G. L. Piggott & H. Goad) as well as by an FCAR doctoral fellowship (to K. Brannen). The authors can be reached at: Dept. of Linguistics, McGill University, 1085 Av. Penfield, Montréal (Qué), Canada, H3A 1A7; e-mail: heather.goad@mcgill.ca, brannen@cyberus.ca

1. Throughout this paper, we use the term 'aspirated' to refer to stops which are *post*-aspirated (cf. Steriade 1997). In this way, aspiration is technically final release (see Laver 1994: 355).

2. For convenience, we will continue to use the term 'coda'.

3. Since both rhymes and nuclei can branch, it may seem that Government Phonology freely permits ternary branching VVC rhymes. In this framework, rhymes of this shape are ruled out by 'strict locality', which requires the head, the initial V in this string, to be adjacent to every other member of the rhyme.

4. Importantly, by focussing on *similar* release properties across types of onsets, we do not mean that the aspiration found on an onset in a CV string and that the release observed on the onset of an empty-headed syllable are formally represented in the same fashion. This will become clear in §§5–6.

5. Jensen points out that word-internal short vowels are sometimes subject to syncope, e.g. /luba – dadu/ → [lubdaːd] 'our breath'. However, it is highly unlikely that the resulting cluster is a coda–onset sequence: the sonority profile is irrelevant to the well-formedness of the resulting cluster; word-internal clusters are otherwise subject to epenthesis, as discussed in the text; and syncope appears to be optional (Thorburn 1993). We concur with Thorburn (1993) that the first consonant in a derived CC cluster is the onset of an empty-headed syllable.

6. The facts of English CVC words are not as straightforward as implied in the text. With the notable exception of [ŋ], the inventory of final consonants permitted in CVC words is very similar to that in CVVC words. This suggests that final consonants in CVC words may also be onsets, a problem which we leave to future research.

7. To our knowledge, such licensing restrictions are observed for laryngeal properties but never for place. As will be seen in §6, this follows from the representations that we provide.

8. The child data in this paper come from the following sources: Mollie (Holmes 1927), Hildegard (Leopold 1939), Jacob (Menn 1978), Lasan (Fey & Gandour 1982), Scott (collected by the first author). All data are transcribed according to the conventions used by the particular author(s) with one exception: final aspiration is represented with a superscript [h] where Leopold uses an apostrophe.

9. In contrast to our view, Fey and Gandour consider the CN̩ sequences to be heterosyllabic, but they provide no evidence for this analysis.

10. On Hildegard's 'Bates', Leopold (1939: 51) remarks that " ... the aspirated [t] [was] added after a short pause ... ". On Jacob's 'box', Menn (p.c., 17 Oct 1999) states: "The hyphen represents an inaudible stretch before the (rather strong) k release ... ".

11. Our focus has been on early obstruent-final CVC forms. One could argue that the timing and release properties observed for such forms are not due to the syllabification of the final consonant as an onset, but are instead due to poor motor control. In Goad and Brannen (2000), we argue against this position by extending the analysis to children's early treatment of nasal-final targets.

12. Some support for our representation comes from Cayuga. Dyck (1990) mentions that in this language, there are two epenthetic elements, aspiration and [e], which are allophones of each other. Aspiration appears in three contexts: t_k, k_t, t_n; while [e] only occurs in one: k_k. Under our analysis, aspiration would arise from 'spreading' the features of the left-most consonant into the intervening nucleus. It would be blocked in favour of [e] in the environment k_k as a violation of the OCP would otherwise result.

13. Avery (1996) analyses English as a 'Contextual Voice' (CV) language, rather than as a 'Laryngeal Voice' (LV) language as we have in (17b). The difference lies in how voiced obstruents are specified. According to Avery, in a CV language, voiced obstruents bear no Laryngeal node at all; they are thus more appropriately labelled as voiceless unaspirated. The facts discussed in this paper do not crucially rely on English being an LV language. What is crucial is that it is not an SV language.

14. (17a) correctly predicts that we consider Lasan's inputs to contrast for SV in both initial and final position. However, initial voiceless targets show the same range of SV-like outputs as do sonorant obstruents, e.g. [^{n}du] 'two'. We propose that this is because voiceless stops share SV with the following vowel in the output. This could be due to SonNode, a constraint which requires every segment to be specified for a 'sonority node', SV or Laryngeal (Kawasaki 1998: 73). SonNode is not satisfied when the voiceless stop is in final position, 'eat' → [i.t^{h}], *[i.dn̩], as sharing of features across a heterosyllabic VC sequence is highly marked cross-linguistically.

15. If SV is genuinely available before Laryngeal, one might wonder why there are not more children like Lasan. We suspect that this is because the bulk of acquisition research has been conducted on English, and VOT values for stops in initial stressed syllables – data which the learner has ready access to – strongly indicate that English is not an SV language. This is in contrast to a language like French where voicing lead on /b, d, g/ may be interpreted by the child as evidence for SV. In view of this, we suggest that the solution to whether the patterns that Lasan exhibits are more widespread than presently documented lies in cross-linguistic acquisition studies.

References

Avery, J. P. (1996). The representation of voicing contrasts. Unpublished doctoral dissertation, University of Toronto.

Blake, B. J., & Dixon, R. M. W. (1979). Introduction. In R. M. W. Dixon & B. J. Blake (Eds.), *Handbook of Australian languages*, Vol I. Amsterdam: John Benjamins.

Blevins, J. (1995). The syllable in phonological theory. In J. Goldsmith (Ed.), *The handbook of phonological theory.* Oxford: Blackwell.

Borowsky, T. (1986). Topics in the lexical phonology of English. Doctoral dissertation, University of Massachusetts, Amherst. Published 1990 by New York: Garland.

Charette, M. (1991). *Conditions on phonological government.* Cambridge: Cambridge University Press.

Cho, Y. (1990). Parameters of consonantal assimilation. Doctoral dissertation, Stanford University. Published 1999 by München: LINCOM Europa.

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht: Foris.

Clements, G. N., & Hume, E. (1995). The internal organization of speech sounds. In J. Goldsmith (Ed.), *The handbook of phonological theory.* Oxford: Blackwell.

Clements, G. N., & Keyser, S. J. (1983). *CV phonology.* Cambridge, MA: MIT Press.

Demuth, K., & Fee, E. J. (1995). Minimal words in early phonological development. Ms., Brown University and Dalhousie University.

Dyck, C. (1990). Cayuga syllable structure and epenthesis. Ms., University of Toronto.

Elson, B. (1947). Sierra Popoluca syllable structure. *International Journal of American Linguistics, 13*, 13–17.

Everett, D., & Seki, L. (1985). Reduplication and CV skeleta in Kamaiurá. *Linguistic Inquiry, 16*, 326–330.

Fey, M., & Gandour, J. (1982). Rule discovery in phonological acquisition. *Journal of Child Language, 9*, 71–81.

Fikkert, P. (1994). On the acquisition of prosodic structure. Doctoral dissertation, Leiden University/HIL. Published 1994 by The Hague: Holland Academic Graphics.

Giegerich, H. (1985). *Metrical phonology and phonological structure.* Cambridge: Cambridge University Press.

Gnanadesikan, A. (1995). Markedness and faithfulness constraints in child phonology. Ms., Rutgers University.

Goad, H. (1997). Codas, word minimality, and empty-headed syllables. *Child Language Research Forum, 28*, 113–122. Stanford: Center for the Study of Language and Information.

Goad, H., & Brannen, K. (1998). *Codas, not! The syllabification of final consonants in early child language*. Paper presented at the Montréal-Ottawa-Toronto Phonology Workshop, University of Ottawa, Feb. 1998.

Goad, H., & Brannen, K. (2000). Syllabification at the right edge of words: Parallels between child and adult grammars. In J. Steele & M. Yoo (Eds.), *McGill Working Papers in Linguistics, 15*, 1–26.

Goldsmith, J. (1990). *Autosegmental and metrical phonology.* Oxford: Blackwell.

Harris, J. (1994). *English sound structure.* Oxford: Blackwell.

Harris, J. (1997). Licensing inheritance: an integrated theory of neutralisation. *Phonology, 14*, 315–370.

Hoard, J. E. (1978). Syllabification in Northwest Indian languages. In A. Bell & J. B. Hooper (Eds.), *Syllables and segments*. Amsterdam: North-Holland.

Holmes, U. T. (1927). The phonology of an English-speaking child. *American Speech, 2*, 219–225.

Ingram, D. (1978). The role of the syllable in phonological development. In A. Bell & J. B. Hooper (Eds.), *Syllables and segments.* Amsterdam: North-Holland.

Itô, J. (1986). Syllable theory in prosodic phonology. Doctoral dissertation, University of Massachusetts, Amherst. Published 1988 by New York: Garland.

Jakobson, R. (1941/1968). *Child language, aphasia, and phonological universals*. The Hague: Mouton.

Jensen, J. (1977). *Yapese reference grammar.* Honolulu: University of Hawaii Press.

Kawasaki, T. (1998). Coda constraints – Optimizing representations. Unpublished doctoral dissertation, McGill University, Montreal.

Kaye, J. (1989). *Phonology: a cognitive view.* Hillsdale, NJ: Lawrence Erlbaum.

Kaye, J. (1990). 'Coda' licensing. *Phonology, 7*, 301–330.

Kaye, J., Lowenstamm, J., & Vergnaud, J.-R. (1990). Constituent structure and government in phonology. *Phonology, 7*, 193–231.

Key, H. (1961). Phonotactics of Cayuvava. *International Journal of American Linguistics, 27*, 143–150.

Ladefoged, P. (1993). *A course in phonetics* (3rd ed.). Orlando: Harcourt Brace Jovanovich.

Laver, J. (1994). *Principles of phonetics.* Cambridge: Cambridge University Press.

Leopold, W. F. (1939). *Speech development of a bilingual child; a linguist's record.* New York: AMS Press.

Lombardi, L. (1991). Laryngeal features and laryngeal neutralization. Doctoral dissertation, University of Massachusetts, Amherst. Published 1994 by New York: Garland.

McCarthy, J., & Prince, A. (1990). Prosodic morphology and templatic morphology. In M. Eid & J. McCarthy (Eds.), *Perspectives on Arabic linguistics II.* Amsterdam: John Benjamins.

Menn, L. (1978). *Pattern, control, and contrast in beginning speech.* Aphasia Research Center: Boston University Medical School.

Mithun, M., & Basri, H. (1986). The phonology of Selayarese. *Oceanic Linguistics, 25*, 210–254.

Piggott, G. L. (1991). Apocope and the licensing of empty-headed syllables. *The Linguistic Review, 8*, 287–318.

Piggott, G. L. (1992). Variability in feature dependency: the case of nasality. *Natural Language & Linguistic Theory, 10*, 33–77.

Piggott, G. L. (1999). At the right edge of words. *The Linguistic Review, 16*, 143–185.

Pinker, S. (1984). *Language learnability and language development.* Harvard: Harvard University Press.

Prince, A., & Smolensky, P. (1993). Optimality theory: constraint interaction in generative grammar. Ms., Rutgers University and University of Colorado, Boulder.

Rice, K., & Avery, P. (1989). On the interaction between sonorancy and voicing. *Toronto Working Papers in Linguistics, 10*, 65–82.

Rice, K. (1992). On deriving sonority. *Phonology, 9*, 61–99.

Rice, K. (1993). A reexamination of the feature [sonorant]. *Language, 69*, 308–344.

Rice, K. (1996). Default variability: the coronal–velar relationship. *Natural Language & Linguistic Theory, 14*, 493–543.

Sapir, J. D. (1965). *A grammar of Diola-Fogny.* Cambridge: Cambridge University Press.

Shvachkin, N. K. (1948/1973). The development of phonemic speech perception in early childhood. In C. Ferguson & D. Slobin (Eds.), *Studies of child language development.* New York: Holt Rinehart Winston.

Smith, N. V. (1973). *The acquisition of phonology: a case study*. Cambridge: Cambridge University Press.

Stafford, R. L. (1967). *An elementary Luo grammar.* Oxford: Oxford University Press.

Stampe, D. (1969). The acquisition of phonetic representation. *Proceedings of CLS, 5*, 443–454.

Stemberger, J. (1996). Syllable structure in English, with emphasis on codas. In B. Bernhardt, J. Gilbert, & D. Ingram (Eds.), *Proceedings of the UBC International Conference on Phonological Acquisition.* Somerville, MA: Cascadilla Press.

Steriade, D. (1982). Greek prosodies and the nature of syllabification. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Steriade, D. (1997). Phonetics in phonology: the case of laryngeal neutralization. Ms., University of California, Los Angeles.

Thorburn, R. (1993). Syllable structure and syllabification in Yapese. *University of Massachusetts Occasional Papers, 16*, 293–318.

Trigo, R. L. (1988). On the phonological derivation and behavior of nasal glides. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Urbanczyk, S. (1996). Patterns of reduplication in Lushootseed. Unpublished doctoral dissertation, University of Massachusetts, Amherst.

Westbury, J., & Keating, P. (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics, 22*, 145–166.

Yoshida, S. (1990). The 'mora' in Japanese. *Phonology, 7*, 331–351.

The phonology-phonetics interface and Syllabic Theory*

Shosuke Haraguchi
Meikai University

0. Introduction

In Haraguchi (1998), I proposed a new theory of syllable structure, which is dubbed the Set Theory of the Syllable (hereafter, STS). This theory is based on the observation that the syllable is the only pronounceable linguistic unit under normal circumstances, as demonstrated by Fujimura (1996a, b).[1,2] It is also based on a number of basic assumptions, summarized in (1)–(4).

(1) In line with Fujimura's Convertor/Distributor (C/D) model (Fujimura 1992, 1994, 1995a, 1996a, b, etc.), ordering of segments is assumed not to be relevant to underlying phonological structure. That is, the underlying phonological structure of syllables consists of unordered sets of underspecified features.

(2) Phonology and phonetics are different in nature in that the former is unordered in the underlying structure, while the latter is ordered.
Linearization applies to unordered phonological syllabic structures from which ordered phonetic (segmental) structures are derived.

(3) A version of underspecification supposed in the C/D model is assumed. (See Fujimura 1996b, to appear.)

I will first outline the STS in Haraguchi (1998), revising the theory slightly and expanding it as well. At the same time, I will focus attention on a number of properties of the phonology-phonetics interface, in particular, a number of linearization processes, both universal and language-particular.

The STS differs from segment-based phonological approaches in that it is syllable-based. In all traditional segment-based approaches, it is assumed that syllable structure is constructed based on existing segmental phonological structure. Consequently, the internal constituents of syllable structure are ordered from the very beginning. This conception is opposite to the STS, which assumes that the internal underlying phonological syllable structure consists of unordered set of distinctive features and that ordered phonetic (segmental) structure is determined by processes of linearization and phonetic implementation.

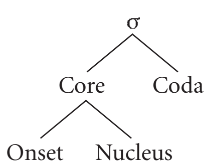
Here, I propose, slightly revising my earlier (1998) system, that a syllable (σ) consists of the syllabic root and peripheries (the p-fix and the s-fix) as indicated in (4a). These are assumed to form an unordered set and the ordering is determined by the system of linearization to be discussed below.

(4) a. σ consists of the sets {{P-fix}, {Syllabic Root}, {S-fix}}.
 b. The syllabic Root consists of the sets {{Onset}, {Nucleus}, {Coda}}.
 c. (i) The sets {Onset}, {Nucleus} form the set {Core}.
  (ii) The sets {Nucleus}, {Coda} form the set {Rime}.

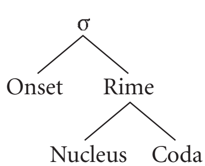
The term 's-fix' ('syllable-suffix'), is virtually equivalent to what Halle and Vergnaud (1980) call Appendix.[3] The term 'p-fix' refers to 'syllable-prefix'. As illustrated in (4b), the syllabic root consists of the set of the onset, the nucleus, and the coda. I propose here that in languages like Japanese, the sets {(Onset)}, {Nucleus} form the set {Core}, as illustrated in (4c.i). On the other hand, in languages like English, the sets {Nucleus}, {Coda} form the set {Rime}, as illustrated in (4c.ii). In other words, I assume that languages are parametrized in this respect.

The syllable structure incorporating (4c.i) is diagrammatically represented as in (5), which is different from the widely-accepted syllable structure incorporating (4c.ii), which is represented as in (6).

(5)



(6)



Notice that the fact that CV is the unmarked syllable structure follows as a consequence of (5): CV forms the core of a syllable and the coda is a periphery of the syllable. Thus, I assume that (5) is an unmarked option, while (6) is a marked one.

The STS incorporates the following system of universal constraints, which I assume constitute part of the linearization algorism

(7) (A) Onset:
Onsets precede the nucleus.
(B) Nucleus:
A nucleus consisting of a vowel (or a sequence of vowels) precedes the coda.
(C) Periphery:
(i) The p-fix precedes the syllabic root.
(ii) The s-fix follows the syllabic root.

These constraints determine the ordering of the onset, the nucleus, the coda, p-fix and s-fix. I assume these to be the principles of linearization, which change phonological set structures into ordered derived structures, which in turn are ultimately turned into temporally-ordered phonetic (segmental) structures.

The phonological codas in Tokyo Japanese are rather simple as will be shown in the next section and they must conform to the following coda condition, first introduced by Itô (1986):

(8) (D) Coda Cond (Japanese):
a. A coda consonant can have either a nasal feature only, or if non-nasal, no place specification of its own at all in the phonological representation.
b. A coda must share the "Place node" with the immediately following onset consonant, if any.

This paper is organized as follows: In Section 1, I review and discuss the syllable structure of Tokyo Japanese and illustrate how the STS works. In Section 2, I review and discuss the syllable in English, which has much more complicated structures than those of Tokyo Japanese. In Section 3, I am concerned with a system of processes of the phonology-phonetics interface. The final section provides a brief summary and conclusion.

1. Tokyo Japanese

1.1 Introduction

The syllable structure of Tokyo Japanese is known to be relatively simple and can be summarized as follows:

(9) (C)V(V)(X)
(where X stands for the moraic nasal or the first half of a geminate.)

This shows that Japanese admits onsetless syllables, with either a short or long nucleus, and it allows either the moraic nasal or the first half of a geminate as the coda.

The examples in (10) illustrate the possible syllable structures in accordance with the structure in (9).

(10) Japanese syllable structures:

i.	V; VV:	a, i, u, e, o; aa, ii, uu, ee, oo, ai, au, etc.
ii.	CV; CVV:	ka, ki, ku, ke, ko; kaa, kii, kuu, kee, koo, etc.
iii.	VN; VVN:	an, in, un, en, on; aan, iin, oon, uun, oon, etc.
iv.	CVN; CVVN:	kan, kin, kun, ken, kon; kaan, kiin, kuun, keen, koon, etc.
v.	VC; VVC:	aQ, iQ, uQ, eQ, oQ; aaQ, iiQ, uuQ, eeQ, ooQ, etc.
vi.	CVC; CVVC:	kaQ, kiQ, kuQ, keQ, koQ; kaaQ, kiiQ, kuuQ, keeQ, kooQ, etc.

(where Q stands for the first half of a geminate.)

1.2 Syllable structures in Tokyo Japanese

Some of the typical examples in (10) are analyzed as in (11):

(11)

i′.	*ai* = {a, i}		cf. *ia		
	au = {a, u}		cf. *ua		
ii′.	*ka* = {k, a}				
	kaa = {k, a, V}	or	{k, a, +long}		
iii′.	*an* = {a} {N}	or	{a, N}		
	aan = {a, V} {nasal}	or	{a, V, nasal}		
iv′.	*kan* = {k, a} {N}	or	{k, a, N}	cf.	*nak, *kna, *akn, *ank
	kaan = {k, a, V} {nasal}	or	{k, a, V, N}	cf.	*naak, *knaa, *aakn, *aank

v′. *aQ* = {a} {oral} or {a, oral}
aaQ = {a, V} {oral} or {a, V, oral}
vi′. *kaQ* = {k, a} {oral} or {k, a, oral}
kaaQ = {k, a, V} {oral} or {k, a, V, oral}
(where Q stands for the first half of a geminate, N, for the moraic nasal, and V, for a long vowel.)

In Japanese, as well as in many other languages, the diphthongs *ai* and *au* in (i′) are the only possible order in a syllable, since no ordering of **ia* and **ua* is permitted syllable-internally.[4] Also 'Q' is phonetically realized as the glottal stop 'ʔ' if no consonant follows; otherwise it is realized as the first half of a geminate, assimilating completely to the following consonant, as illustrated in (12).

(12) *kappa* = {k, a, Q}+{p, a} (where Q = oral consonant or X slot.)
kaC pa
Root

The moraic nasal 'N' is realized as a nasal consonant with the same Place specification as the following consonant by assuming assimilation as in (13a).

(13) a. *kanda* = {k, a, N}+{d, a}
kaN da
Place

b. k a n
a

If no consonant comes after the moraic nasal 'N', it assimilates to the preceding vowel, as illustrated in (13b).

Take, for example, the syllable structures of *ka* and *kaa* in (11ii′), which can be analyzed as follows:

(14)

Lexical Item	Formal Rep.	Abbreviatory Rep.	
a. *ka* (mosquito)	= {dorsal, +low}	= {k, a}	cf. *ak
b. *kaa* (car)	= {dorsal, +low, V}	= {k, a, V}	cf. *aak

The word *ka* in (10a) is analyzed as the set {dorsal, +low} in its formal representation, or informally as the set {k, a}. Basically, I will use the abbreviatory notation wherever a more formal notation is not needed. Note that Japanese has no word like **ak*, due to the coda condition in (8).

The structures of *an* 'plan' and *na* 'vegetable' can be distinguished by representing each of them as follows:

(15) a. *an* : σ = {a} {N}
b. *na* : σ = {n, a}

Another way of distinguishing (15a) and (15b) is to assume that {N} consists of just the set {nasal}, with no Place specification, and that {n} is the set {nasal, apical}, with the Place specification 'apical'. Assuming in line with Coda Condition in (8) that only the former is permitted to appear in the coda in this case, the two words in question can also be represented as follows:

(16) a. *an* : σ = {a, N} = {a, nasal}
b. *na* : σ = {n, a} = {a, {nasal, apical}}[5]

Similarly, the following words, which are slightly more complicated, can be represented in either way, as in column I or column II.

(17)

		I		II	
a.	*kan* :	σ = {k, a} {N}	or	{k, a, N}	cf. *nak, *kna, *akn, *ank
b.	*kaan* :	σ = {k, a, V} {nasal}	or	{k, a, V, N}	cf. *naak, *knaa, *aakn, *aank
c.	*aQ* :	σ = {a} {oral}	or	{a, oral}	
d.	*aaQ* :	σ = {a, V} {oral}	or	{a, V, oral}	
e.	*kaQ* :	σ = {k, a} {oral}	or	{k, a, oral}	
f.	*kaaQ* :	σ = {k, a, V} {oral}	or	{k, a, V, oral}	

(where Q stands for the first half of a geminate.)

For example, the word *kan* can be analyzed as in I or as in II. This is because Tokyo Japanese does not permit words such as **nak*, **kna*, **akn*, **ank*, etc. If there is no convincing evidence to the contrary, the simpler internal representation should be chosen, which means (16) is preferred to (15) and the analyses of (17a–f) in II are preferred to the ones in I.

Notice thus that the syllable structure in column I of (18) corresponds to (19a) whereas that of column II corresponds to (19b).

(18)

	I		II
a.	*nin*	vs.	*ni*
b.	*nan*	vs.	*na*
c.	*nun*	vs.	*nu*
d.	*nen*	vs.	*ne*
e.	*non*	vs.	*no*

(19) a. {n, V, N}
b. {n, V}
(where V stands for either {*i*, *a*, *u*, *e*, *o*}.)

Japanese exhibits a somewhat more complicated syllable structure in front of -*ko* "inhabitant" (and possibly other restricted morphemes).

(20) a. *Rondonkko* 'Londoner' CVN-CVNQ-CV
b. *Wiinkko* 'Viennese' CVVNQ-CV
c. *Tookyookko* 'Tokyoite' CVV-CVVQ-CV
(where N stands for the moraic nasal, Q for the first half of a geminate.)

The first half of these geminates (represented by Q) can be derived by insertion of an X node in front of -*ko*, and by the subsequent assimilation process, which spreads the root of the velar *k* to the X node, thus making the basic syllable structure of Tokyo Japanese slightly simpler.

1.3 Inflectional forms of Japanese Verbs

Let us turn to inflectional forms of Japanese verbs. Bloch (1946a, b) classifies Japanese verbs into two types: C(onsonant)-ending and V(owel)-ending. Consider *kak* 'write' and *tabe* 'eat', which are representative examples:

(21) a. C-ending verbs: *kak* 'write'
b. V-ending verbs: *tabe* 'eat'

Some inflectional forms of these two types of verbs are as in (22) and (23).

		Consonant-ending	Vowel-ending
(22)	Verb stem	*kak-*	*tabe-*
	Negative	*kak-a-nai*	*tabe-nai*
	Preverbal	*kak-i-mas-u*	*tabe-mas-u*
	Prenominal	*kak-u*	*tabe-ru*
	Present	*kak-u*	*tabe-ru*
	Provisional	*kak-eba*	*tabe-reba*
	Tentative	*kak-oo*	*tabe-yoo*
	Imperative	*kak-e*	*tabe-ro*
	Past	*ka-i-ta*	*tabe-ta*
(23)	Passive	*kak-are-ru*	*tabe-rare-ru*
	Capable	*kek-e-ru*	*tabe-ra-re-ru; tabe-re-ru*
	Causative	*kak-as-u*	*tabe-sas-u*

Note that the stem-final consonant /k/ is deleted before a theme vowel /i/ + the Past morpheme /ta/ in (22). The morphological structure of a word is responsible for the determination of the relative ordering of the verbal roots and suffixes.

If we assume that when a consonant-ending verb is followed by the negative morpheme *nai*, a theme vowel /a/ is inserted, then the underlying phonological form should be (24i). Assuming, on the other hand, that the negative morpheme has two allomorphs *nai* and *a-nai*, then the consonant-ending verb selects a negative allomorph *a-nai*, the underlying phonological form should be (24ii), with subsequent incorporation of /k/ as the onset of the syllable /a/.

(24)	*kak-anai*	(i)	{{k, a}}. {{k }}. {{n, a, i}}.	+ *a*-Insertion
		(ii)	{{k, a}}. {k} {{a}}. {{n, a, i}}.	+ *k* Onset
		(where the period (.) indicates a syllable boundary.)		

To put it differently, in (i), the syllable concatenator /a/ becomes necessary, while, in (ii), we need to introduce two allomorphs, *nai* and *a-nai*. The question is which of these two is correct.

To answer this question, let us focus attention to (the theme vowels and) the affixal morphemes in (23) and (24) and compare the following two allomorphs:

(25)		C-ending Verb	V-ending Verb
	Negative	*a-nai*	*nai*
	Preverbal	*i-mas-u*	*mas-u*
	Prenominal/Present	*u*	*ru*
	Provisional	*eba*	*reba*
	Tentative	*oo*	*yoo*
	Imperative	*e*	*ro*
	Past	*i-ta*	*ta*
	Passive	*are-ru*	*rare-ru*
	Capable	*e-ru*	*(ra-)re-ru*
	Causative	*as-u*	*sas-u*

If we examine these allomorphs with care, we notice that we cannot predict what vowels or what consonants are selected, admitting the existence of some subregularities. For example, the Imperative allomorphs /e/ and /ro/ should be lexically introduced in the lexicon. Thus, we must introduce two allomorphs for each of these morphemes.

Interestingly, all of these allomorphs have the effect of avoiding violations of the Obligatory Contour Principle (OCP). Thus, all the allomorphs adjoined to C-ending verbs begin with a vowel, while those adjoined to V-ending verbs begin with a consonant. This clearly shows that so-called OCP effects are observable even in the lexicon.

So far, I have discussed the outline of my new theory of syllable structure, showing how the STS works based on several syllabic properties of Japanese. In the next section, I will survey how this new theory is applicable to account for various syllabic phenomena of English, whose syllable structure is much more complicated than that of Japanese.

2. Syllable structure of English

2.1 Introduction

Let us now turn to the syllable structure of English. As is well known, English has a much more complicated range of possible syllables than Japanese. The syllable structure of English can be approximately represented as in (26).

(26)

		t			N		{t/p/k}		
	s	p	l	V(V)	l	s	p	th	s
		k	r		r		k	s	
								t/d	
								p/b	
								k/g	

(where N stands for *n, m, ŋ*.)

As Fujimura (1996b) points out, English has no contrast between /sp, st, sk/ and /*sb, *sd, *sg/ nor does it have any contrast between /sp, st, sk/ and /*ps. *ts, *ks/ in syllable-initial position. In syllable-final position, on the other hand, it has a contrast between /sp, st, sk/ and /ps, ts, ks/, as illustrated in (27)–(29).

(27) a. *task* /tæsk/ /sk/ : {dorsal, spirantized}
b. *tax* /tæk.s/[6] (Fujimura 1996b)

(28) a. *cast* /kæst/ /st/ : {apical, spirantized}
b. *cats* /kæt.s/[6]

(29) a. *lasp* /læsp/ /sp/ : {labial, spirantized}
b. *lapse* /læp.s/[6]

Note that the dot (.) in these examples indicates that the word-final /s/ is an s-fix. Here, following Fujimura, I assume /sp, st, sk/ are single units: spirantized labial, apical, and dorsal, respectively. As will be discussed in detail below, he analyzes for example /sk/ as forming a unit {dorsal, spirantized} while /ks/ is analyzed as a sequence of the coda {dorsal} and the s-fix {s}.

Note here that the word task /tæsk/ cannot be analyzed as in (30).

(30) {dorsal, spirantized, æ, apical}

This is because this structure could alternatively be linearized also as in (31).

(31) a. *skat* /skæt/

In this connection, (30) should not be linearized as in (32), because the word-final /t/ and /s/ in (32) are s-fixes.

(32) a. *asked* /æsk.t/
b. *tax* /tæk.s/

I have assumed that English selects (4c.ii) as its basic syllable structure. (4c) is repeated here for ease of reference.

(33) c. (i) The sets {Onset}, {Nucleus} form the set {Core}.
(ii) The sets {Nucleus}, {Coda} form the set {Rime}.

Under (33), the words *asked* and *tax* in (32) should be analyzed as (34a) and (34b) respectively.

(34) a. {{æ, sk}}. $\{t\}_{\text{s-fix}}$
b. {{t}, {æ, k}}. $\{s\}_{\text{s-fix}}$

Recall also that the s-fixes {t} and {s} are placed to the right of the coda by the linearization principle in (7C), which is repeated here for ease of reference, and other principles of linearization.

(7) (C) Periphery:
(i) The p-fix precedes the syllabic root.
(ii) The s-fix follows the syllabic root.

As shown in Halle and Vergnaud (1980: 95) and Haraguchi (1998), the s-fixes in English and German are restricted to apicals due to the following constraint:

(35) Constraint on the S-fix:
In the unmarked case, only apicals are permitted as the s-fix.

Without assuming this constraint both in English and in German, we could not account for the fact that s-fixes are restricted to apicals in these languages (see Halle & Vergnaud 1980: 95 for discussion of German cases).

2.2 On the status of /th/ in English

Let us now review and discuss some controversial cases in English, i.e., the status of the nominal suffix /θ/. Before discussing its status, let us first mention briefly that English permits at most three consonant clusters in both the onset and the coda. Note that the nominal suffix -*th* /θ/ and the final /s/ of the four consonant clusters like /nkθs/ observed in word-final position of *strengths* are s-fixes and thus they cannot constitute the coda. Haraguchi (1998) assumes, following Fujimura (1996a: 62) and Yip (1991), that English and possibly other languages as well have the following constraints:

(36) Place Constraint:
Onsets and codas in English can have one Place specification.

This constraint excludes consonant clusters in the onset and/or the coda.

(37) a. *tk, *dg, *tp, *db, etc.
b. *tm, *km, *png, etc.

To illustrate how the STS analyzes the English word *strengths*, consider the representations in (38).

(38) a. Morphological Structure: $[[[\text{strɛng}]_{\text{A}}\ \text{th}]_{\text{NP}}\ \text{s}]_{\text{Pl}}$
b. Phonological Structure: $\{\text{st, r, ɛ, Ng}\}\ \{\theta\}_{\text{s-fix}}\ \{\text{s}\}_{\text{s-fix}}$
b′. {spirantized, apical,
rhotacized, -high, -low,
-back, nasal, dorsal} $\{\text{interdental, -voiced}\}_{\text{s-fix}}\ \{\text{fricative}\}_{\text{s-fix}}$

In (28), I assume that $[\text{strɛng}]_{\text{A}}$ is a positional allomorph of $[\text{strɔng}]_{\text{A}}$. (38a) is the morphological structure of *strengths* and (38b) is an abbreviatory notation for (38b′). As for the legitimacy of this structure of /strɛNg/, see the discussion in Section 2.3. It should be clear that the ordering of {th} and {s} relative to {strenk} is determined by the morphological structure, as well as the principle in (7C.ii). The linearization of /st, r, ɛ/ and /Ng/ is determined both by some version of the sonority hierarchy and by a number of linearization principles particular to English to be discussed below. As pointed out by Osamu Fujimura, and briefly discussed above, there is no contrast between *sp/st/sk* and **sb/*sd/*sg* in the onset and the coda. However, there is a contrast between the sequences *sp/st/sk* and *ps/ts/ks* in the coda, but there is no such a contrast in the onset.

Let us now consider some of the processes of phonetics-phonology interface. Linearization of {spirantized, apical, rhotacized, -high, -low, -back, nasal, dorsal} is determined by the following processes, which will be formalized later:

(E) a. The features {nasal, dorsal} (i.e., {Ng}) are assigned to the coda by a language-particular process in English.
b. Other consonantal features are assigned to the onset of the syllable and vocalic features are assigned to the nucleus.
c. The feature [spirantized] is assigned to onset-initial position.
d. The ordering of other features is determined by a version of the Sonority Hierarchy.

(Ea) assigns {Ng} to the coda. I will return to (Ea) in Section 2.3 below. The other processes in (E) determine the phonetic ordering of /strɛ/. Note that, as mentioned above, (Ec) follows from the fact that English does not permit /*ts, *ps, *ks/ word-initially as well as syllable-initially.

The example in (38) is interesting and important in that it is a clear counterexample to Fujimura's claim in (39).

(39) Voicing Principle Regulating the S-fix:
"The phonetic voicing status [of s-fixes] must agree with that of the coda."
(Fujimura, to appear: 3)

As shown in Haraguchi (1998), the nominal suffix [θ] should be phonologically specified as [-voiced], with which the preceding coda velar [g] assimilates and not vice versa. It is also noteworthy that assimilation is obligatory in the case of velar but optional in the case of dental clusters, as illustrated in (40).

(40) a. *broad* *breadth* [bred/tθ]
b. *wide* *width* [wid/tθ]

The nominal suffix [θ] optionally turns the dental coda into the corresponding voiceless segment. These examples show that the "s-fix" phonologically specified as [-voiced] remains as is, and (optionally) turns the obstruent coda into [-voiced] by assimilation. Notice that Regressive Voicing Assimilation here is identical to that which applies to phrases as illustrated in (41).

(41) *have to* /hæv tə/ > [hæftə]

In this connection, consider the case of *warmth*, cited in Fujimura (1996b) and Haraguchi (1998):

(42) Morphological Structure: [[[warm]$_A$ th]$_N$

Phonological Structure: {{w}, {a, r, m}}. {θ}$_{s\text{-fix}}$

As pointed out in Haraguchi (1998), this is a clear counterexample to the principle in (39). Fujimura is thus forced to analyze {θ} of *warmth* as the coda. This is contradictory in that it violates the constraint in (36). Furthermore, it is unnatural and inconsistent since /θ/ is analyzed as the s-fix in (38), while it is analyzed as the coda in (42).

Notice that the ordering of 's-fixes' and 'p-fixes' with respect to the coda and the onset is determined by the interaction of (4) and the linearization processes in (7C), repeated again for reference.

(4) a. σ consists of the sets {{P-fix}, {Syllabic Root}, {(S-fix}}.
b. Syllabic Root consists of the sets {{(Onset)}, {Nucleus}, {Coda}}.

(7) (C) Periphery:
(i) The p-fix precedes the syllabic root.
(ii) The s-fix follows the syllabic root.

The ordering principles in (7C) can most likely be subsumed as a part of the definitions of 'p-fix' and 's-fix, respectively.

Consider now the word *sixths*, which has the morphological structure represented in (43a) (Haraguchi 1998).

(43) *sixths*:
a. Morphological Structure: [[[six]$_A$ th]$_N$ s]$_{Pl}$
b. Phonological Structure: {{s}, {{i}, {k}}$_{Coda}$.{s}$_{s\text{-fix}}$.{θ}$_{s\text{-fix}}$.{s}$_{s\text{-fix}}$}

Since the coda has only one Place specification, /s/ to the right of /k/ should be analyzed as an s-fix. This shows that the suffixes in (43), i.e., /th/ and /s/, should be analyzed as s-fixes. This observation again shows that we are led to the conclusion that /th/ of (42) should be analyzed as an s-fix.

Based on the observations so far, we can conclude that the nominal suffix *th* should be lexically specified as [-voiced] and that this suffix is not a coda but an s-fix.

2.3 Some merits of the STS

Let us now overview a number of merits of the present theory, based on Haraguchi (1998). First of all, consider example (44), which is pointed out in Cairns and Feinstein (1982) and Cairns (1988):

(44) hip = {h, i, p} cf. *pih, *iph, *ihp (Cairn & Feinstein 1982)

The word *hip* needs no information on the relative ordering of /h/, /i/ and /p/. This is so because /h/ can be only in the onset, the vowel /i/, in the nucleus, and the remaining consonant /p/, in the coda.

No segment-based syllabic theories can account for this fact because such traditional theories construct syllable structure based on the ordered segmental structure. In the present theory, on the other hand, assuming the unordered set structure of distinctive features, (44) can be accounted for without any additional ad hoc mechanisms. Note that the STS can account not only for this example but also for numerous similar facts exemplified in (45). It should be clear that no specification of ordering of these cases is necessary:

(45)	a.	*hob*	=	{h, a, b}	*bah, *abh, *ahb
	b.	*hit*	=	{h, i, t}	*tih, *ith, *iht
	c.	*hitch*	=	{h, i, tʃ}	*tchih, *itchh, *ihtch
	d.	*hid*	=	{h, i, d}	*dih, *idh, *ihd
	e.	*hic*	=	{h, i, k}	*kih, *ikh, *ihk
	f.	*hiss*	=	{h, i, s}	*sih, *ish, *ihs
	g.	*his*	=	{h, i, z}	*zih, *izh, *ihz
	h.	*him*	=	{h, i, m}	*mih, *imh, *ihm
	i.	*hike*	=	{h, a, i, k}	*kaih, *aikh, *aihk
	j.	*heat*	=	{h, i, V, t}	*tiih, *iith, *iiht
	k.	*hat*	=	{h, æ, t}	*tah, *ath, *aht
	l.	*hound*	=	{h, a, u, n, d}	*ndauh, *aundh, *auhnd
	m.	*hive*	=	{h, a, i, v}	*vaih, *aivh, *aihv
	n.	*hole*	=	{h, o, u, l}	*loh, *olh, *ohl
	o.	*hone*	=	{h, o, u, n}	*noh, *onh, *ohn
	p.	*hoof*	=	{h, u, V, f}	*fooh, *oofh, *oohf
	q.	*heel*	=	{h, i, V, l}	*leeh, *eelh, *eehl

In English, and possibly in many other languages, /h/ exclusively occupies the onset, whereas the other consonants in these cases should be assigned to the coda. Thus, the process of linearization forces /h/ to appear in the onset, /i/ and other vowels in the nucleus, and the consonants /p, t, k, etc./ in the coda.

This linearization can be determined either by the negative principle in (46) or by the positive principle in (47) (Haraguchi 1998).

(46) [h] does not occur in the coda.

(47) [h] exclusively occupies the onset.

Since (47) entails (46), these cannot be regarded as simply notational variants. At present, however, I have no convincing argument for choosing one of these two principles over the other.

Whichever turns out to be correct, however, it should be clear that the STS, unlike previous segment-based syllabic theories, can account for the above examples including /h/.

Also, note that, as shown in Haraguchi (1998), these are not isolated examples. There are a number of other cases in which the ordering of segments can be left unspecified in the underlying phonological structure. Consider the following examples containing /-Ng/:

(48)	a.	*sing*	=	{s, i, Ng}	*ngis, *ings, *isng
	b.	*king*	=	{k, i, Ng}	*ngik, *ingk, *ikng
	c.	*ping*	=	{p, i, Ng}	*ngip, *ingp, *ipng
	d.	*ting*	=	{t, i, Ng}	*ngit, *ingt, *itng
	e.	*ding*	=	{d, i, Ng}	*ngid, *ingd, *idng
	f.	*ring*	=	{r, i, Ng}	*ngir, *ingr, *irng
	g.	*ling*	=	{l, i, Ng}	*ngil, *ingl, *ilng
	h.	*wing*	=	{w, i, Ng}	*ngiw, *ingw, *iwng
	i.	*binge*	=	{b, i, Ndʒ}	*ngeiw, *ingeb, *ibnge
	j.	*thing*	=	{θ, i, Ng } etc.	*ngith, *ingth, *ithng

(49)	a.	*song*	=	{s, ɔ, Ng}
	b.	*long*	=	{l, ɔ, Ng}
	c.	*Kong*	=	{k, ɔ, Ng}
	d.	*sang*	=	{s, æ, Ng}
	e.	*gang*	=	{g, æ, Ng}
	f.	*sung*	=	{s, ʌ, Ng}
	g.	*cling*	=	{k, l, i, Ng}
	h.	*string*	=	{s, t, r, i, Ng}
	i.	*spring*	=	{s, p, r, i, Ng} etc.

Haraguchi (1998) points out that no specification of the ordering in the examples of (48) is necessary. The process of linearization in (50) forces /Ng/ to appear in the coda, the vowel /i/ in the nucleus, and the other consonants, such as /p, t, k/, in the onset in (48). Similarly, in (49), the velar nasal /Ng/ also occupies the coda, vowels the nucleus, and other consonants the onset. The principle in (50), which bans the word-initial cluster *Ng*, and other relevant principles of linearization can guarantee the ordered phonetic linear structures.

(50) /Ng/ is prohibited to occur in the onset.

Now consider the examples similar to (48) and (49).

(51)	a.	*think*	=	{th, i, Nk}	*nkith, *inkth, *ithnk
	b.	*sunk*	=	{s, ʌ, Nk}	*nkus, *unks, *usnk
	c.	*kink*	=	{k, i, Nk}	*nkuk, *unkk, *uknk

In these examples as well, no specification of the ordering is required. /Nk/ should appear in the coda and other consonants in the onset, and the vowels /i, u/ in the nucleus.

Note that the set {Ng} is directly realized phonetically as velar nasal [ŋ], as shown in (52).

(52) {N, dorsal, voiced} → [ŋ]

This means that no deletion of [g] is necessary in word-final position. The set {Nk}, on the other hand, is realized as the ordered phonetic structure [nk] by a process of linearization and assimilation.

The STS can account for other cases in which no information concerning ordering is phonologically necessary, such as cases where the onset and the coda have the same consonant. To illustrate this, observe the examples in (53).

(53)	a.	*babe* :	{b, e, i, b}	k.	*pup* :	{p, ʌ, p}
	b.	*cack* :	{k, æ, k}	l.	*sass* :	{s, æ, s}
	c.	*dad* :	{d, æ, d}	m.	*souse* :	{s, a, u, s}
	d.	*dead* :	{d, e, d}	n.	*rear* :	{r, i, r}
	e.	*gag* :	{g, æ, g}	o.	*deed* :	{d, i, V, d}
	f.	*judge* :	{dʒ, ʌ, dʒ}	p.	*peep* :	{p, i, V, p}
	g.	*kick* :	{k, i, k}	q.	*pipe* :	{p, a, i, p}
	h.	*lull* :	{l, ʌ, l}	r.	*cake* :	{k, e, i, k}
	i.	*mom* :	{m, a, m}	s.	*tight* :	{t, a, i, t}
	j.	*nun* :	{n, ʌ, n}	t.	*mime* :	{m, a, i, m}

Haraguchi (1998) argues that if two identical consonants appear in a syllable, one necessarily occupies the onset and the other the coda. This is due to the fact that there is no geminate permitted in English. Thus, it should be clear that we have no need to specify the ordering of the units in the sets illustrated above.

Consider now the word *state*, which contains a spirantized apical /st/ and an apical /t/. The syllable structure of this word cannot be analyzed as in (54):

(54) *state* = {st, e, i, t}

This is because English has a word such as *taste*. Thus, the word *state* should have the following syllable structure:

(55) {{st}, {e, i, t}}

In this section, I have discussed three cases in which information on the ordering of elements is not necessary in the underlying structure. The observation above will be enough to show the legitimacy of the STS.

2.4 Where does linearization apply?

Let us now turn to consider the question of what level the processes of linearization should apply at. Haraguchi (1998) argues that there are a number of cases in which ordered phonological structure is required. In fact, (56) shows that we must apply linearization word-internally. Insertion of [ɪ] requires that the word-final sibilant *s* and the plural suffix [S] be adjacent to each other. Similarly, the word-final [t, d] and the Past morpheme (as well as the Past participial morpheme) [D] stand adjacent to each other.

(56) a. *bus* *buses*
b. *base* *bases*

(57) a. *state* *stated*
b. *respect* *respected*

Whether we assume that [ɪ] is inserted or that the allomorphs /iS/ and /iD/ are selected,[7] linearization should have applied internal to these words (Haraguchi 1998). This is so because [s] and [t] must be in word-final position if the Obligatory Contour Principle (OCP) forces [ɪ] to be inserted or the allomorphs /iS]/ and /iD/ to be selected in these plural and Past (Participle) forms. This observation shows that linearization should first apply word-internally.

Assimilation requires adjacency of the two elements. This shows that linearly ordered structure must be obtained before the application of assimilation. Adjacency of two elements is also necessary in the case of Velar Softening in (58):

(58) a. *critic* /k/ *criticize* /s/
b. *fungus* /g/ *fungi* /dʒ/

At the level where the velars /k, g/ undergo Velar Softening, the velars and /i/ should be located adjacent to each other. Thus, linearization should apply before the application of the process in question.

The observation above has the important theoretical consequence that ordering of segments is necessary at some point of phonological structure. This contradicts Fujimura's claim to the effect that ordering is entirely a matter of phonetic implementation.

2.5 On the internal structure of the syllable

Let us now discuss the internal structure of the syllable. Given the choices in (4c), which is repeated here for ease of reference, I have assumed that English selects (4c.ii).

(4) c. (i) The sets {Onset}, {Nucleus} form the set {Core}.
 (ii) The sets {Nucleus}, {Coda} form the set {Rime}.

This selection leads to the choice of (59b) out of the following two logically possible analyses of cat:

(59) a. *cat* = {{k, æ}, {t}}
 b. *cat* = {{k}, {æ, t}}

Note that *cat* cannot be represented simply as the set {k, æ, t}, since this set can be linearized as either *tack* [tæk] or *cat* [kæt].

There are several pieces of evidence suggesting the correctness of (4c.ii). Let us review my (1998) arguments. As Takeru Honma (personal communication) has pointed out, English has a number of restrictions concerning the combination of the nucleus and the coda, while there is virtually no restriction concerning the combination of the onset and the nucleus (also Bernhardt & Stemberger 1998).

Take for example the diphthong *au*, which can occur only before a dental or a dental cluster as illustrated in (60). It does not occur before a labial or a velar, which is supported by the fact that English has no words, such as **oup, *ouk*. etc. as indicated in (60).

(60) a. *out* [aut] *oup, *ouk
 b. *loud* [laud] *loup, *louk
 c. *bound* [baund] *boump, *bounk

A cooccurrence restriction of this kind should be stated as a condition which operates within the rime. This and other facts indicate that (4c.ii) is appropriate in English.

In Haraguchi (1998), I also argue in favor of (4c.ii) based on the well-known secret languages or disguised languages in English, which show the necessity of a split between the onset and the rime. Two such secret languages are Pig Latin and the *op (ob)*-language, described in (61) and (62) respectively.

(61) a. Pig Latin: Move the initial consonant sequences in the word, if any, to the end, and then add the sequence [*ēy*] to the right.
(Chomsky & Halle 1968: 342–343)

b. Onset – X → X – Onset + *ēy*

(62) a. Insert [op] (or [ob]) between every onset and rime of a word.
(Haraguchi 1982: 205–208)

b. Onset – Rime → Onset – *op* – Rime

Thus, the words *Pig Latin* and *secret*, for example, become as in (63) and (64) respectively.

(63) *Pig Latin*
a. Pig Latin: [ɪgpēy ætənlēy]
b. *Op* Language: [popɪg lopætopən]

(64) *secret*
a. Pig Latin: [iːkrɪtsēy]
b. *Op* Language: [sopiːkropɪt]

Notice that [op] is not inserted in other positions; e.g., between the nucleus and the coda. These secret languages can be best accounted for if we assume that there is a split between the onset and the rime.

Taking these kinds of facts into consideration, Haraguchi (1998) concludes that there be a break between the onset and the rime. For further evidence to this effect, see Bernhardt and Stemberger (1998: 440–441) among others.

Assuming the rime structure, we can formally represent the difference between *on* and *no* in English, as in (65).

(65) a. *on* : {{o, n}}
b. *no* : {{n}, {o}}[8]

Otherwise, these two words could not be distinguished phonologically. Notice that this contrast in English is rather different from what we have seen in (16a) and (16b) in Japanese in Section 1.1.

2.6 On the importance of morphological structure

Haraguchi (1998) argues that morphological structure is extremely important for the determination of syllable structure and linearization. By assuming a certain morphological structure, we can easily account for the phonological structure of the word *warmth*, as illustrated in (66).

(66) Morphological Structure: [[[warm]$_A$ th]$_{NP}$

Phonological Structure: {{w}, {a, r, m}}. {θ}$_{s\text{-}fix}$

As Fujimura (1996b: 62) and Yip (1991) note, codas and onsets can have only one Place specification in English. Thus, the bilabial /m/ should be the only one which appears in the coda in this case. Thus, the interdental or coronal /θ/ should be an s-fix. As already discussed above, this contradicts Fujimura's claim, in (39), that the voicing status of s-fixes always agrees with that of the coda.

2.7 The Voicing Contour Principle and its implications

Another noteworthy observation by Fujimura concerns the relation of syllable structure and voicelessness:

(67) The relation of syllable structure and voicelessness:
If the margin is obstruent and not specified to be voiced, then there has to be a voicing discontinuity taking place at some point within the core [=Syllabic Root —SH], and there is no further change of voicing within the same syllable. If there is no obstruent element in either side of the nucleus, then the syllable must be voiced throughout the syllable on that side. (Fujimura 1996b: 72)

This is no doubt an important observation, for it is tantamount to saying that there is no voiceless island within a syllable. In fact, voiceless sequences appear between the syllable boundary and a voiceless obstruent in English. Fujimura calls this the Voicing Contour Principle, which can be stated as in (68).

(68) The Voicing Contour Principle:
If there is a voiceless obstruent in the onset or the coda of a syllable, the voicelessness should continue (or spread) to the syllable-initial or syllable-final boundary, respectively. (Haraguchi 1998)

This principle accounts for Voicing Assimilation of the Plural morpheme -*s*, the Present tense morpheme -*s*, and the Past (Participle) morpheme -*ed* in English.

(69)	a.	Sg.	Pl.	b.	Sg.	Pl.
		node	*nodes*		*note*	*notes*
		cab	*cabs*		*cap*	*caps*
		dog	*dogs*		*pack*	*packs*

(70)	a.	noun	Genitive	b.	noun	Genitive
		boy	*boy's*		*cat*	*cat's*
		cab	*cab's*		*cap*	*cap's*
		bag	*bag's*		*pack*	*pack's*

(71)	a.	Present	Past	b.	Present	Past
		decide	*decided*		*kiss*	*kissed*
		dive	*dived*		*camp*	*camped*
		shrug	*shrugged*		*bank*	*banked*

These morphemes are rather exceptional or marked in that they are the only ones that undergo progressive assimilation. Notice that all other assimilations are regressive in English. If we assume that these apical morphemes have no phonological specification of voicing in the lexicon, their assimilatory behavior automatically follows from the Voicing Contour Principle in (68).

Notice also that the Voicing Contour Principle accounts for the disyllabic nature of the examples in (72):

(72) a. *little, bottle, cycle, sample, goggle*
b. *center, enter, glitter*
c. *kitten, mutton, cotton*

The liquids [l, r] and the nasal [n] are regarded as syllabic in these cases, which is guaranteed by this principle as well as the sonority principle, which says that more sonorous segments come before less sonorous segments in the coda.

It should be pointed out that there are a number of languages in which this principle in (68) does not apply. Consider, for example, the following Georgian examples:

(73)		Nominative	Genitive		
	a.	*mts'erali*	*mts'erl-is*	'writer'	
	b.	*mt^{h}vrali*	*mt^{h}vral-is*	'drunk person'	
	c.	*mt^{h}argmneli*	*mt^{h}argmnel-is*	'translator'	(Chitoran 1998)

The word-initial bilabial nasal /m/, which clearly constitutes a "p-fix" (i.e., syllable prefix), does not conform to the Voicing Contour Principle in (68). This suggests that this principle either is not selected in Georgian or is ranked so low that its effects are not observable in this language.

Another example comes from Swedish, due to Olle Kjellin (personal communication):

(74) *optimism* 'optimism'

The word-final bilabial nasal /m/, which can be optionally devoiced, remains voiced even after the voiceless consonant /s/. This also violates the Voicing Contour Principle in (68). Again, we can say that this principle either is not selected in Swedish or is ranked very low.

Notice incidentally that in both cases, the consonant violating the Voicing Contour Principle in (68) is the bilabial nasal /m/. An interesting question is whether this is merely accidental or whether there is a principled reason for it. The answer appears to be that it is accidental. For example, Icelandic permits a case like (75), as pointed out by Anderson (1999) .

(75) *pü:kr* 'concealment'

This example also shows, however, that Icelandic does not conform to the principle in question.

Based on the observations above, we can conclude that: (i) this constraint, if it is correct as is, is either parametrized or ranked very low or (ii) the validity of Voicing Contour Principle in (68) is questionable and thus the principle requires serious reconsideration.

3. On the system of the phonology-phonetics interface

Let us now turn to some processes governing the phonology-phonetics interface. In Section 2.2, I suggested the following processes of linearization, repeated here for ease of reference.

(E) a. The features {nasal, dorsal} (i.e., {Ng}) are assigned to the coda by a language-particular process in English.
  b. Other consonantal features are assigned to the onset of the syllable and vocalic features are assigned to the nucleus.
  c. The feature [spirantized] is assigned to onset-initial position.
  d. The ordering of other features is determined by a version of the Sonority Hierarchy.

In this section, I discuss a number of further universal principles and language-particular ones.

Let us begin by considering some of language-particular principles. Recall that the ordering of the set {Ng} in English is determined by the following principle:

(76) (=50) /Ng/ is prohibited to occur in the onset.

Due to the interaction of (76) and (77), the ordering of {Ng} is determined; it can only be located in the coda.

(77) (=35) Constraint on the S-fix:
In the unmarked case, only apicals are permitted as the s-fix.

Recall also that the ordering of the set {h} is determined either by (46) or (47), repeated here as (78) and (79) respectively.

(78) (=46) [h] does not occur in the coda.

(79) (=47) [h] exclusively occupies the onset.

One of these will constitute a linearization principle of English, and possibly of other languages.

The principles of linearization should also incorporate (80).

(80) a. In the unmarked case, the vocalic features occupy the nucleus only.
b. In English, a sonorant can occupy the nucleus in some marked cases.

Note that (80a) is universal whereas (80b) is a marked principle of English and possibly some other languages.

If the set contains two identical consonants, recall that one of them is assigned to the onset and the other to the coda. To guarantee this, we need the principles of linearization in (81) and (82).

(81) If the set contains two identical consonants, assign one of them to the coda.

(82) All the remaining consonants are assigned to the unoccupied onset or to the unoccupied coda.

(80) and (81) determine the linear order of two identical consonants as well as any other remaining consonants.

Consider now the cases of spirantized stops. Since spirantized stops can occur either in the onset or in the coda, (Ec) should be specified more generally, as in (83).

(83) The feature [spirantized] is ordered before the stop consonantal features.

If the onset and/or the coda consist of two or more consonants, the ordering of these consonants is determined by the following linearization principle, based on the so-called sonority hierarchy:

(84) The Sonority Principle of Linearization:
Consonants in a syllabic root are ordered depending on the sonority hier-

archy: Less sonorous consonants are positioned farther from the syllable nucleus and more sonorous consonants closer to the syallble nucleus.

This constraint determines the ordering of the word-initial consonants of the word *strong*, as illustrated in (85).

(85) {st, r, ɔ, Ng} : [str ɔ ŋ]

The relative order of [t] and [r] is determined by the sonority principle in (84).

It is interesting to note that (84) guarantees that in the syllabic root of every language the degree of sonority is lowered depending on the relative peripherality of consonant; i.e., if they are located more peripheral position, they become less sonorous.

Though there are no doubt many more principles of linearization, both universal and language-particular, I think these principles suffice for the determination of the ordering of Japanese and English. These principles would seem to constitute part of the core of the phonology-phonetics interface.

4. Concluding remarks

I have argued that the STS opens up the possibility for a new conception of phonological structure. Also it can account for a number of facts that were either unexplained or unnoticed in the previous theories. I have shown that differences in the basic syllable structures of Japanese and English follow partly along the lines of (4c.i) and (4c.ii).

(4) c. (i) The sets {Onset}, {Nucleus} form the set {Core}.
(ii) The sets {Nucleus}, {Coda} form the set {Rime}.

This theory also fits in well with the basic tenets of the C/D model of syllable structure, proposed in Fujimura's pioneering papers.

From this new perspective, things look very different from the familiar segment-based theory of syllable structure. The STS makes it possible to reexamine previously accepted conceptions of syllable structure and the entirety of sound structure of natural languages. I have also presented a number of crucial cases that can best be accounted for by the STS.

It should be added that the STS opens up a new possible way of looking at, not only phonological structure, but also the the phonology-phonetics interface. It should also help to reveal hidden or unnoticed regularities.

One might wonder how the unordered underlying phonological structure is acquired based on ordered phonetic forms. Notice to begin with it is assumed that information on the ordering of syntactic units is irrelevant to syntax. I assume that the situation in phonology is parallel to that of syntax. Information on ordering is not essential to language. It is enforced by the phonetic temporal constraint on linguistic units. Thus, the human brain has the capacity to mentally represent only the essential properties of language, based on the temporally-enforced input information.

Finally, I must admit there remain a number of challenges for the present theory to meet, which can be summarizd as follows:

i. to make the system of linearization much more explicit than the above.
ii. to examine syllable structures of a variety of languages whose syllable structures are more complicated than those of Japanese and English.
iii. to make the STS more explanatorily adequate by the detailed and profound examination of the universal principles and language-particular ones.
iv. to shed further light on the essential nature of the syllable.

Notes

* This paper is an expanded and partly revised version of Haraguchi (1998). This paper is in part supported by the Monbusho under Grant-in-Aid for Scientific Research (A) No. 07401015, and under Grant-in-Aid for COE Research No. 08CE1001, and Special Research Project of the Investigation of Languages and Cultures of the East and the West of the University of Tsukuba. I am grateful to Osamu Fujimura, for long discussions on the C/D Model and to Junko Itô and Armin Mester for giving me a chance to present an earlier version of this paper at a colloquium at the University of California, Santa Cruz, where I received a number of valuable comments and suggestions from the audience. My thanks also go to Roger Martin, Thomas Purnell, Jeroen van de Weijer, John Shillaw and John Whitman for many helpful comments and suggestions on an earlier draft of this paper.

1. A similar conception is also found in pre-generative phonological theories. For example, Haugen (1956: 216) proposes "that the syllable be defined as the smallest unit of recurrent phonemic sequences". See also Halle and Vergnaud (1980: 95).

2. Fujimura assumes that the syllable is a unit of utterance. I assume that it is also a unit of language in general.

3. Halle and Vergnaud (1980: 95) propose this concept based on Moulton's analysis of consonant clusters of Modern German. They show that the appendix in German can have at most three segments, all of which consist of coronal nonsonorant consonants.

4. Adam Sherman of the University of California at Santa Cruz has pointed out to me that Hebrew admits /ia/ as a diphthong. This suggests that languages are parametrized as to whether the vowel sequence [ia] constitutes a diphthong.
Note incidentally that the diphthong [au] is rather rare in native Japanese words, though it is sometimes found in loan words, as illustrated in (i)(see Katayama 1998):

(i) a. *auto* 'out'
b. *sutauto* 'stout'
c. *kauti* 'couch'
d. *saundo* 'sound'
e. *maundo* 'mound'

5. A parallel observation also applies to the following pairs in Japanese:

a. *an, in, un, en*
b. *na, ni, nu, ne*

6. In (27)–(29), the dot (.) indicates that the immediately following consonant /t/ is an s-fix.

7. Note that if we assume the allomorphs /iS/ and /iD/ are chosen to be correct, these allomorphs have the set structures {i} {S}} and {{i} {D}}.

8. I assume that the vowel is short at the lexical level, but undergoes word-final lengthening, as shown by Chomsky and Halle (1968).

Appendix

Consonantal Feature (English)

place/manner	dorsal D	coronal C	apical A	labial L	no place
stop	**k**	**ch/jh**	**t**	**p**	
fricative		**sh/zh**	**s/z**	**f/v**	
spirantized	**sk**		**st**	**sp**	
interdental			**th/dh**		
nasal	**ng**		**n**	**m**	
lateral					**l**
rhotacized					**r**
palatalized					**y**
labiovelarized					**w**
glottalized					**h/Φ**

Symbols in the table cells are pseudo-orthographic.
Cited from Fujimura (1998) and from his lecture notes.

References

Anderson, Stephen R. (1999). The Nature of Derivational Relations in Morphology. Paper presented at the Colloquium of Tokyo Metropolitan University.

Bernhardt, Barbara Handford, & Joseph Paul Stemberger (1998). *Handbook of Phonological Development: From the Perspective of Constraint-Based Nonlinear Phonology*. San Diego: Academic Press.

Bloch, Bernard (1946a). Studies in Colloquial Japanese Part I, Inflection. *Journals of the American Oriental Society, 66*.

Bloch, Bernard (1946b). Studies in Colloquial Japanese Part III, Derivation of the Inflected Words. *Journals of the American Oriental Society, 66*.

Cairns, Charles E. (1988). Phonotactics, Markedness and Lexical Representation. *Phonology, 5*(2), 209–236.

Cairns, Charles E., & Mark H. Feinstein (1982). Markedness and the Theory of Syllable Structure. *Linguistic Inquiry, 13*(2), 193–225.

Chitoran, Ioana (1998). *Georgian Harmonic Clusters: Phonetic Cues to Phonological Representation*. Paper presented at the meetings of LP'98. Published in *Phonology*, X, cc–ccc.

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

Fujimura, Osamu (1989). Demisyllables as Sets of Features: Comments on Clements' Paper. In J. Kingston & Mary E. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech* (pp. 334–340). Cambridge: Cambridge University Press.

Fujimura, Osamu (1992). Phonology and Phonetics – A Syllable-based Model of Articulatory Organization. *Journal of Acoustic Society of Japan (E), 13*, 39–48.

Fujimura, Osamu (1994). The C/D Model: A Computational Model of Phonetic Implementation. In E. Ristad (Ed.), *Language Computations (DIMACS Series 17)* (pp. 1–20). Providence: American Mathematical Society.

Fujimura, Osamu (1995a). Prosodic Organization of Speech Based on Syllables: The C/D Model. *Proceedings ICPhS 95. 3*, 10–17.

Fujimura, Osamu (1995b). The syllable: Its Internal Structure and Role in Prosodic Organization. In B. Palek (Ed.), *Proceedings of LP '94: Item Order in (Natural) Languages* (pp. 53–93). Prague: Charles University.

Fujimura, Osamu (1996a). *The C/D Model as a Dynamic, Non-segmental Approach (ATR Technical Report, TR-H-184)*. Kyoto: ATR Human Information Processing Research Laboratories.

Fujimura, Osamu (1996b). Syllable Structure Constraints: A C/D Model Perspective. *SWOT Proceedings: Working Papers of U. C. Irvine*, 59–74.

Fujimura, Osamu (to appear). Syllable Feature and the Temporal Structure of Speech. Ed. by Bohumil Palek *Proceedings of LP'96*.

Halle, Morris, & Jean-Roger Vergnaud (1980). Three Dimensional Phonology. *Journal of Linguistics Research, 1*(1), 83–105.

Halle, Morris, & Jean-Roger Vergnaud (1987). *An Essay on Stress*. Cambridge, MA: MIT Press.

Haraguchi, Shosuke (1982). *Kotoba no Bunka* (Languages and Cultures). Tokyo: Kobian Shobo.

Haraguchi, Shosuke (1991). *A Theory of Stress and Accent*. Dordrecht: Foris.

Haraguchi, Shosuke (1998). *A Theory of the Syllable*. Paper presented at LP'98, Columbus, Ohio, September 15, 1998.

Haugen, Einar (1956). The Syllable in Linguistic Description. In Morris Halle et al. (Eds.), *For Roman Jakobson* (pp. 213–221). The Hague: Mouton.

Itô, Junko (1986). Syllable Theory in Prosodic Phonology. Doctoral dissertation, University of Massachusetts, Amherst [Published in 1988 by Garland].

Katayama, Motoko (1998). Markedness of [au] in Japanese: Evidence from Loan-words. Paper read at Phonology Forum 1998 in Kobe.

Kubozono, Haruo (1993). *The Organization of Japanese Prosody*. Tokyo: Kuroshio.

Levin, Juliette (1985). A Metrical Theory of Syllabification. Doctoral dissertation, MIT.

McCarthy, John, & Alan Prince (1993). Prosodic Morphology I: Constraint Interaction and Satisfaction. Ms., University of Massachusetts, Amherst, and Rutgers University.

McCarthy, John, & Alan Prince (1995). *Faithfulness and Reduplicative Identity*. University of Massachusetts, Amherst & Rutgers University.

Prince, Alan, & Paul Smolensky (1993). Optimality Theory: Constraint Interaction in Generative Grammar. Ms., Rutgers University and University of Colorado, Boulder.

Steriade, Donca (1982). Greek Prosodies and the Nature of Syllabification. Doctoral dissertation, MIT.

Yip, Moira (1991). Coronals, Consonant Clusters, and the Coda Condition. *Phonetics and Phonology, 2*, 61–78.

Hungarian as a strict CV language*

Krisztina Polgárdi
University of Szeged

0. Introduction

In this paper, I propose that Hungarian is best analysed as a 'strict CV' language in the sense of Lowenstamm (1996); that is, as a language where syllables are built up of a single consonant followed by a single vowel. Such an analysis is rather abstract. However, I will argue that significant generalisations are missed, unless this level of abstractness is employed.

I use two main arguments to support the above claim. The first is that there are hardly any phonotactic restrictions between members of (word-internal) consonant clusters in Hungarian. This has generally been used as an argument in Government Phonology (GP) to separate such consonants by an inaudible nucleus. The second argument comes from the behaviour of long vowels, namely, that they are prohibited from preceding consonant clusters, regardless of which consonants constitute these clusters. To be able to account for this restriction in a uniform way, and to be able to provide an explanation for it, all clusters in Hungarian need to be broken up by an empty nucleus.

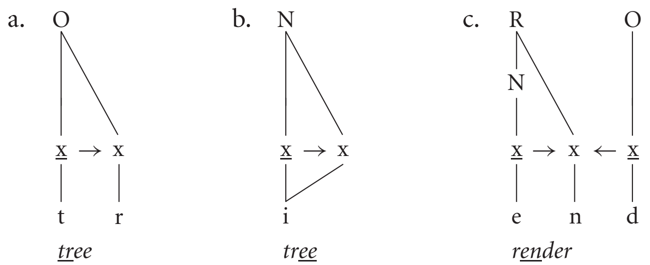
Two of the long vowels, however, *á* and *é*, display aberrant behaviour in this respect. Namely, they can freely occur in front of consonant clusters. These vowels differ from the other long vowels of the system in other respects as well, since they take part in length alternations that the other long vowels do not (cf. the rule of Low vowel lengthening, and the phenomenon of Shortening stems). I will argue that all three peculiarities of these vowels stem from their representation which differs from that of the other long vowels. More precisely, I propose to represent *á* and *é* as sequences of two short vowels.

The data used in this paper are mostly based on Törkenczy (1994) and Nádasdy and Siptár (1994), supplemented by my own research, mostly utilising the reverse dictionary of Papp (1969). In section 1, only monomorphemic forms are considered. Note that all examples in this paper are given in Hungarian orthography.[1]

1. Standard Government Phonology and Hungarian syllable structure

The framework of Government Phonology offers a highly restrictive theory of syllable structure (cf. Kaye, Lowenstamm, & Vergnaud 1990). In this theory, three types of syllabic constituents are recognised, the O(nset), the N(ucleus) and the R(hyme), which are defined by the notion of 'government'. Each syllabic constituent can be simple or branching. The branching structures are illustrated in (1). In the following representations, head positions are underlined, and government is indicated by the arrows '→' and '←'.

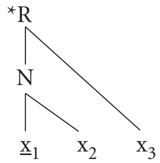
(1) Syllabic constituents



A branching onset (in (1a)) comprises a consonant cluster, such as in the English word *tree*. There are severe restrictions on which pairs of consonants can co-occur in such a structure. A branching nucleus can be of two types. Either, it consists of a long vowel, as the /i:/ in *tree*, in (1b), or it consists of a diphthong. Finally, a branching rhyme, as in (1c), consists of a vowel followed by a coda consonant. There is, however, a further restriction on this case, namely, that a branching rhyme must be supported by a following consonant which is less sonorous than the coda consonant within the rhyme. The underlined portion of the English word *render* constitutes such a well-formed branching rhyme. There is thus a strong relation (a 'governing' relation) between the two consonants in configurations like (1a) and (1c).

Because government is defined as a strictly local and strictly directional relationship, all syllabic constituents can be maximally binary. In this way, long vowels in closed syllables are ruled out, as illustrated in (2), because x_1, the head, is not adjacent to x_3, and thus cannot govern it.

(2) No superheavy rhymes



However, it seems that Hungarian defies this condition, because we can find superheavy rhymes both in word-final and in word-internal position, as illustrated in (3).

(3)

a. word-finally:		b. word-internally:	
– VVCC #		– VVCCV –	
áld	'bless'	férfi	'man'
ráng	'jerk'	bálna	'whale'
férc	'tack'	bástya	'bastion'
észt	'Estonian'	némber	'wench'

What is interesting about these examples is that only the vowels *á* and *é* can occur in such superheavy rhymes (with the exception of *tószt* 'toast' and about a dozen cases word-internally, mostly containing the vowel *ó*) (cf. also Törkenczy 1994: 343–345). I will come back to this generalisation later on.

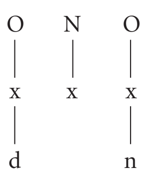
However, this is not the only strange property of Hungarian syllable structure. The inventory of syllabic constituents in (1) predicts that in a well-behaved language only such consonant clusters can be found that can constitute either a complex onset, as in (1a), or a coda-onset cluster, as in (1c). Since the mirror image of a coda-onset cluster does not always result in a well-formed complex onset, certain combinations will be ruled out. This is the situation in English. For example, the word-internal coda-onset cluster *-nd-* is well-formed, but its mirror image *-dn-* is ruled out, since it cannot form a complex onset. Hungarian, however, behaves differently. There are very few restrictions on what consonants can form a possible word-internal cluster. (One such restriction imposes homorganicity on nasal plus stop clusters. Another prohibits clusters where both members involve either a fricative or an affricate.) Thus in Hungarian both *-nd-* and *-dn-* are well-formed, and many other pairs, as illustrated in (4). (The source of these data is Törkenczy 1994: 361.)

(4) Word-internal consonant clusters

a. *Coda-Onset*			b. *Complex Onset*		
lp	alpári	'vulgar'	**pl**	paplan	'quilt'
lf	csalfa	'false'	**fl**	kifli	'crescent-roll'
rk	szarka	'magpie'	**kr**	bokréta	'bouquet'
rd	erdő	'forest'	**dr**	nadrág	'trousers'
jb	lajbi	'waistcoat'	**bj**	gereblye	'rake'
			c. *Onset-Onset*		
szk	deszka	'board'	**ksz**	buksza	'purse'
sp	ispán	'bailiff'	**ps**	ipse	'bloke'
mcs	csámcsog	'champ'	**csm**	kocsma	'pub'
nd	bendő	'belly'	**dn**	bodnár	'cooper'
nk	lankad	'flag'	**kn**	akna	'mine'
nf	fanfár	'fanfare'	**fn**	sufni	'lean-to'
nyty	kulipi[n^y]tyó	'shanty'	**tyny**	sa[t^y]nya	'puny'
rm	lárma	'noise'	**mr**	kamra	'pantry'
kt	akta	'file'	**tk**	patkó	'horseshoe'
c. *Onset-Onset*					
kf	bakfis	'teenage girl'	**fk**	cafka	'whore'
gz	lagzi	'wedding'	**zg**	mézga	'gum'
csk	macska	'cat'	**kcs**	bakcsó	'heron-type bird'

The table in (4) contains two main columns, where examples on the left containing an internal C_1C_2-cluster correspond to examples on the right containing an internal cluster of the opposite order, i.e. C_2C_1. Examples under (4a) could in principle constitute well-formed coda-onset clusters and be represented as (1c). Likewise, examples under (4b) could form perfect complex onsets and be represented as (1a). However, examples under (4c) do not comply with the requirements of either of these structures. Such clusters have been represented in Government Phonology as two consecutive onsets separated by an empty nucleus, as shown in (5).

(5) Onset-Onset, "bogus cluster"



Consonants in such configurations are not strictly adjacent to each other and have therefore received the name "bogus clusters". The most important difference between (5) on the one hand and (1a) and (1c) on the other hand is that in (5) there is no governing relation between the consonants, whereas in (1a) and (1c) that is the defining property of the clusters.

Such an analysis would give us three possible types of consonantal configurations in Hungarian: complex onsets, coda-onset clusters and bogus clusters. However, already in early versions of Government Phonology, if in a language members of apparent clusters do not show phonotactic restrictions, this fact is used as an argument to analyse all clusters in that language as bogus (cf. Kaye, Lowenstamm, & Vergnaud 1990). Here I will argue that Hungarian should be analysed in this manner as well.

An additional argument for a 'bogus cluster only' analysis is provided by the data in (6).

(6) only *á* and *é* preceding internal consonant clusters

a. *Coda-Onset*			b. *Complex Onset*		
lk	nélkül	'without'	**kl**	cékla	'beetroot'
rf	férfi	'man'	**fr**	páfrány	'fern'
rt	Márta	(name)	**tr**	nátrium	'sodium'
			c. *Onset-Onset*		
mzs	kámzsa	'cowl'	**zsm**	pézsma	'musk'
lt	táltos	'shaman'	**tl**	átlag	'average'
lm	Kálmán	(name)	**ml**	sámli	'footstool'
ng	csángó	'Hungarian native in Moldavia'	**gn**	Ágnes	(name)

In (3), we have already seen that the only long vowels that can occur in closed syllables in Hungarian are the vowels *á* and *é*. If the examples in (4) were indeed manifestations of three different structures, then we would expect that the restriction on long vowels would only apply to the case of closed syllables, that is, to the case of coda-onset clusters, illustrated in (6a). On the other hand, we would expect long vowels to be freely allowed preceding complex onsets or onset-onset clusters. This expectation, however, is not borne out, as shown by (6b) and (6c). As can be seen from these examples, long vowels are restricted to *á* and *é* preceding any type of consonant clusters.[2]

A standard GP analysis could be saved by arguing that there are no complex onsets in Hungarian. (Arguments for this claim have been provided by Törkenczy and Siptár (1999).) Under such an analysis, complex onsets would be re-analysed as bogus clusters, and the context for "shortening", or length restriction, would be defined by a disjunction: it would apply in closed syllables and preceding an empty nucleus. The latter context has been identified as a triggering environment for shortening in languages like Turkish and Yawelmani by Kaye (1990). However, this analysis is still unsatisfactory.

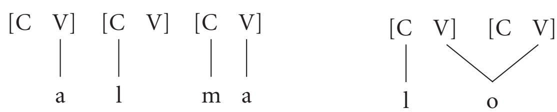
The only way to express the restriction on long vowels in a unitary way is to analyse *all* clusters in Hungarian as spurious. The context for the length restriction then will be defined by a following empty nucleus (the second context mentioned above). I will explain the motivation for this shortly. These facts thus call for a 'strict CV' analysis of Hungarian, following Lowenstamm (1996), which I will introduce briefly in the next section.

2. A strict CV analysis

In this recent version of Government Phonology, it is claimed that languages universally only allow CV-type syllables. This means that closed syllables and long vowels are now analysed as sequences of two CV-syllables, with certain components of syllable structure remaining empty. This is illustrated in (7). In (7a), a word like *alma* 'apple' in Hungarian, with an initial closed syllable, is now re-analysed as a sequence of three CV-syllables.

(7) Strict CV (Lowenstamm 1996)

a. *closed syllable* b. *long vowel*



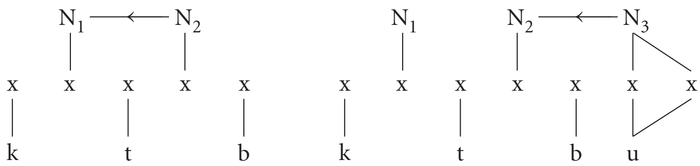
alma 'apple' *ló* 'horse'

A long vowel, on the other hand, which has previously been analysed as a branching nucleus in (1b), is now re-analysed as a sequence of two CV-syllables, with an empty onset intervening. This is illustrated in (7b), with the Hungarian word *ló* 'horse', containing a final long vowel.

The appearance of such empty structure, however, is not unrestricted. Most importantly, sequences of empty nuclear positions are not permitted (*ØCØ). This restriction is motivated in Government Phonology by the phenomenon of vowel ~ zero alternations. An example is provided from Moroccan Arabic in (8) (cf. Kaye, Lowenstamm, & Vergnaud 1990). This phenomenon is accounted for by the notion of Proper Government and the Empty Category Principle, given in (9a–b).

(8) Moroccan Arabic vowel ~ zero alternation

a. *Singular* b. *Plural*



ktɨb ‘I write’ kitbu: ‘we write’

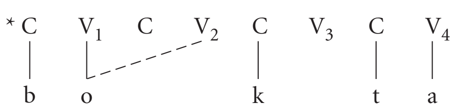
(9) a. *Proper Government* (Kaye 1990)
A nuclear position A properly governs a nuclear position B iff
(i) A governs B (adjacent on its projection) from right to left
(ii) A is not properly governed
(iii) there is no intervening governing domain

b. *Empty Category Principle* (ECP)
An empty Nucleus is phonetically interpreted iff it is not properly governed.

The alternating vowels in (8) are regarded as originally empty. These will remain inaudible as long as they are followed (properly governed) by a full nucleus. N_2 in (8a) thus, not being properly governed, will be pronounced. N_2 can now properly govern N_1, which remains silent. In (8b), on the other hand, the plural marker *-u*: can properly govern N_2, whose pronunciation becomes unnecessary. N_1, however, is now not properly governed, and as a result, it has to be realised phonetically. In this way, an alternating pattern of full and empty nuclei is derived.

Now we can return to the story of Hungarian. The data in (4) and (6) have pointed to a strict CV analysis of Hungarian syllable structure. As I have argued, the only way to express the fact that word-internal clusters behave in a unitary way in restricting the length of a preceding vowel is to analyse all such clusters as bogus. What we still need to explain is why the presence of an empty nucleus inhibits the preceding vowel from being long. To see this, consider the representation of a hypothetical Hungarian form in (10).

(10) impossible long vowel



In this theory, a long vowel is lexically only specified in its head position (that is, in its leftmost position), and it is subsequently spread to its dependent position, indicated by the dashed line in (10). Lowenstamm (1996) proposes that the target of spreading must be licensed, as given in (11).

(11) The target of spreading must be licensed. (Lowenstamm 1996)

Proper government is one way of licensing an empty position. What is important here is that if V_2 was part of a branching nucleus, proper government would have no access to it. The joint effect of the strict CV representation in (10) and the principle in (11) is that long vowels are only allowed if they are followed by a single consonant and a full vowel. (Although the principle in (11) is a stipulation in its present form, note that it is not completely ad hoc, since it applies to geminates as well, prohibiting them from preceding another consonant.)

This concludes the discussion on why Hungarian is best analysed as a strict CV language. One reason is that there are very few restrictions on what type of consonants can co-occur in a word-internal cluster. Another reason is that long vowels are prohibited in front of all types of clusters. However, a problem remains with the long vowels *á* and *é*, since – as was demonstrated in (3) and (6) above – these do occur before consonant clusters. I turn to this problem in the next section.

3. The problem of long *á* and *é*

What is interesting about the long vowels *á* and *é* is that it is exactly these two in the whole system that also differ qualitatively from their short counterparts, namely long /a:/ corresponds to short /ɔ/ and long /e:/ to short /ε/; whereas for all the other vowels the difference is purely quantitative. The vowel system is given in (12).

(12) Hungarian vowel inventory

short vowels			long vowels		
i	ü	u	í	ű	ú
	ö	o	é	ő	ó
e[ε]		a[ɔ]			
					á

One solution could be to propose that *á* and *é* in this context are really short phonologically, and the difference from *a* and *e* is purely qualitative. However, there are a number of arguments against this solution. One argument concerns the symmetry of the vowel system, exhibited in (12). If *á* and *é* were really short vowels, the system would contain nine short vowels and only five long ones. More seriously, in those dialects which distinguish a close *ë* [e] from an open *e* [ɛ], the opposition between a short *ë* [e] and a long *é* [e:] would become difficult to represent.

Another argument is provided by the definite-indefinite conjugation paradigms in Hungarian, illustrated in (13).

(13) Definite-indefinite conjugation

	Indefinite	*Definite*	
a.	vonna	vonná	‘draw COND’
	fenne	fenné	‘whet COND’
b.	vont	vonta	‘draw PAST’
	fent	fente	‘whet PAST’

In the conditional forms (13a), the indefinite-definite distinction is expressed by the *a*~*á*/*e*~*é* difference. If we compare these forms with the past tense forms in (13b), we can see that the difference in definiteness is really expressed by the absence versus presence of *a/e*. It seems thus that two adjacent *a*’s are realised as an *á* and two adjacent *e*’s as an *é*. This provides an argument for representing the *a/á* and *e/é* distinctions as a matter of length, rather than as a matter of quality.[3]

Finally, there are two phonological alternations in Hungarian affecting the vowel pairs *a*~*á* and *e*~*é* which can be argued to involve the dimension of length instead of a qualitative distinction. The first one is the so-called Low Vowel Lengthening rule, illustrated in (14) (cf. Nádasdy & Siptár 1994: 67–70; Rebrus 2000).

(14) Low Vowel Lengthening (LVL)

a.	fa	fát	‘tree NOM ~ ACC’
	létra	létrák	‘ladder SG ~ PL’
	kutya	kutyául	‘dog NOM ~ ESS-MOD’
	marha	marhába	‘cattle NOM ~ ILL’
	lusta	lustább	‘lazy ~ COMPAR’
	alma	almánként	‘apple NOM ~ DISTR’
b.	kefe	kefét	‘brush NOM ~ ACC’
	medve	medvék	‘bear SG ~ PL’
	mérce	mércéül	‘measure NOM ~ ESS-MOD’

vese	vesébe	'kidney nom ~ ill'
fekete	feketébb	'black ~ compar'
körte	körténként	'pear nom ~ distr'

This rule lengthens word-final *a* and *e* when they come to precede a suffix. As can be seen in the examples in (14), the phonological shape of the suffix is irrelevant for the process. It only fails to apply in front of a number of suffixes which could be called "super-analytic" (the term originates from Rebrus 2000: 911), and which show a clitic-like behaviour. Otherwise the process is fully productive.

The other alternation involves the so-called shortening stems, illustrated in (15) (cf. Nádasdy & Siptár 1994: 71–76).

(15) Shortening stems

a.	nyár	nyarak	'summer sg ~ pl'
	madár	madarat	'bird nom ~ acc'
	szamár	szamara	'donkey nom ~ 3sg poss'[4]
	sár	saras	'mud ~ muddy'
	pohár	poharanként	'glass nom ~ distr'
	bogár	bogaracska	'bug ~ dim'
b.	kéz	kezek	'hand sg ~ pl'
	tél	telet	'winter nom ~ acc'
	tenyér	tenyere	'palm nom ~ 3sg poss'
	jég	jeges	'ice ~ icy'
	levél	levelenként	'leaf nom ~ distr'
	veréb	verebecske	'sparrow ~ dim'

This alternation affects a closed class of stems (about 70 items), most of which contain an alternating *á* or *é* (about a dozen contain an alternating high vowel, such as *víz* ~ *vizet* 'water nom ~ acc', which I will disregard here). Only synthetic suffixes, in the terminology of Kaye (1995), trigger the alternation. These are those suffixes that are included in an unstructured phonological domain together with the stem (corresponding to the category of Level 1 suffixes in Lexical Phonology). These suffixes also trigger the processes of *v*-Augmentation and Vowel ~ zero alternation (cf. Rebrus & Polgárdi 1997).[5]

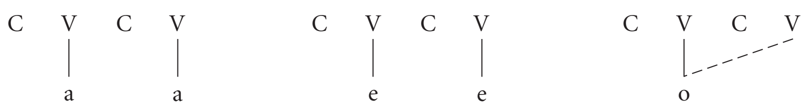
The problem with an analysis treating the pairs *a*~*á* and *e*~*é* as differing only in terms of quality is that this difference is not the same for the two pairs. Long *á* is lower than short *a*, whereas long *é* is higher than short *e*. In an analysis that posits only qualitative differences between the members of these pairs, it thus becomes very difficult to express the changes illustrated in (14) and (15) in a unified manner.

What is interesting about the length-alternations in (14) and (15) is that (apart from the exceptions mentioned above) they only involve the vowels *á* and *é*, and none of the other long vowels shown in (12) – and there are no other length-alternations of any productivity in the language.[6] *á* and *é*, in turn, are precisely those long vowels in the system that are unexpectedly allowed to occur preceding a consonant cluster. In my view, these facts are not accidental. In what follows, I will derive all three generalisations from the representation of these long vowels which will differ from the representation of other long vowels in the system.

More precisely, I propose that the peculiarity of *á* and *é* lies in their representation, given in (16a–b); that is, in the fact that they consist of a sequence of short vowels phonologically, whereas all the other long vowels have the usual representation, illustrated by the example of *ó* in (16c).

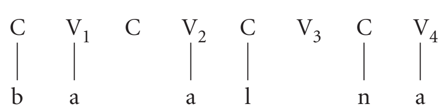
(16) Representation of *á* and *é*

a. /a:/ b. /e:/ c. /o:/



A word like (17), with a long *á* preceding a consonant cluster, will then be well-formed, because it does not violate any conditions. (Note that these representations of *á* and *é* do violate the OCP. However, OCP violations sometimes are, and sometimes are not repaired in different languages. Therefore, I will not deal with this problem any further here.)

(17) possible "long" vowel preceding a consonant cluster



bálna

So far these representations might seem rather ad hoc. But in fact, they predict that *á* and *é* should take part in the alternations illustrated in (14) and (15), whereas the other long vowels should not. To be able to show this, I need to introduce some additional facts of Hungarian.

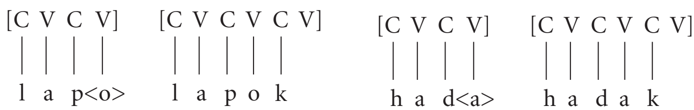
3.1 Low vowel lengthening

Before turning to an analysis of Low vowel lengthening, first let us look at (18).

(18) Stems ending in non-high short vowels

a. "normal" stems	SG	PL		b. "lowering" stems	SG	PL	
o	lap	lapok	'sheet'	**a**	had	hadak	'army'
ö	tök	tökök	'pumpkin'	**e**	fül	fülek	'ear'
e	géz	gézek	'lint'				

Rebrus and Polgárdi (1997):

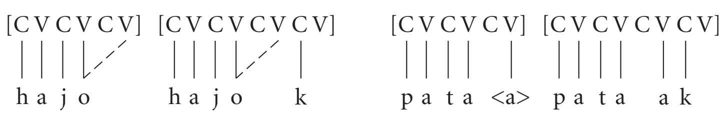


The data in (18) have been generally analysed as involving epenthesis after the stem and before consonant-initial suffixes. "Normal" stems, in (18a), take a (harmonising) mid vowel after them, whereas the lexically specified class of "lowering" stems, in (18b), takes a (harmonising) low vowel as epenthetic.[7] In Rebrus and Polgárdi (1997), we argued that the underlined vowels in (18) are not epenthetic; neither do they belong to the following suffix. Instead, they must be specified at the end of the stem. In front of synthetic suffixes, these short non-high vowels surface, but they have to remain silent, when they occur in domain-final position (indicated by angled brackets in (18)). This means that under this view every stem in Hungarian ends in a vowel. (Note that the final empty nucleus in *lapok* and *hadak* is licensed by the domain-final parameter in Hungarian (cf. Kaye 1990).)

Now compare these data with those in (19).

(19) Stems ending in non-high long vowels

a. "normal" stems	SG	PL		b. LVL stems	SG	PL	
ó	hajó	hajók	'ship'	**á**	pata	paták	'hoof'
ő	cipő	cipők	'shoe'	**é**	kefe	kefék	'brush'



Stem-final long mid vowels, in (19a), do not show length alternation. (19b), on the other hand, contains stems manifesting Low vowel lengthening. If we analyse the examples in (19b) as cases of domain-final shortening, instead of lengthening preceding a suffix,[8] then an interesting parallelism with (18) arises. As we have seen, in (18) final short non-high vowels underwent underparsing. If long *á* and *é* are represented as sequences of two short vowels, the final short non-high vowels of (19b) will automatically be affected by the same process of underparsing as they were in (18). Stems in (19a), on the other hand, end in a long vowel which does not satisfy the structural description of this process, and therefore here shortening will not apply. The difference between the behaviour of final long mid and low vowels thus falls out automatically from their representational differences which are independently motivated.

Looking at the representation of *pata* in (19b), the question might arise why (after underparsing) *a* does not spread to the final position, resulting in a long *á* after all. This can be prevented by referring to the principle of Structure Preservation, according to which no new segment types can be created in the lexicon (cf. Kiparsky 1985). Since a representation such as (16c) does not exist for *á* (and *é*), the final vowel of forms like *pata* will remain short and cannot re-lengthen.

Note that the representation of *hajók* in (19a), as it stands, violates the principle formulated in (11). While in the singular the second V-position of *ó* is licensed by virtue of being domain-final, the same V-position in the plural is no longer at the end of the domain, instead it is followed by another empty nucleus. The issue is more general, since this configuration does not only arise as a result of suffixation, but it is also present in monomorphemic forms (e.g. *kór* 'disease'). The same phenomenon can be found in other languages as well. For the case of Norwegian, Lowenstamm (1996) proposes that final empty nuclei in that language have the same licensing potential as full vowels do and are thus strong enough to license spreading into a preceding V-position. Here I will adopt the same solution for Hungarian.

Note also that there is a slight difference in the domain of application of underparsing in (18) versus (19). This is illustrated in (20).

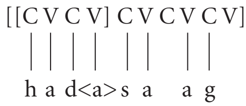
(20) Different domains of underparsing

a. analytic suffixation

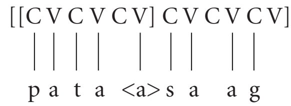
```
[[C V C V] C V C V]          [[C V C V C V] C V C V]
  | | | |  | | |               | | | |   |  | | |
  h a d<a>b a n                p a t a   a  b a n
```

hadban 'army INESS' *patában* 'hoof INESS'

b. super-analytic suffixation



hadság ‘army + N-forming’



pataság ‘hoof + N-forming’

The difference between (20a) and (20b) shows that shortening of domain-final long low vowels only applies before super-analytic suffixes, as in *pataság*, but it does not in front of analytic suffixes, as in *patában*. Underparsing of domain-final short non-high vowels, on the other hand, applies before all analytic domain boundaries, that is, in front of all analytic suffixes; and we get *hadban* and *hadság*. This difference in application will have to be specified independently.

A remaining problem involves that group of words which end in a non-alternating *á* or *é*, illustrated in (21).

(21) Non-alternating final *á* and *é*

a.	(Feri) bá	(Feri) bát	‘uncle (Frank) NOM ~ ACC’
	burzsoá	burzsoát	‘bourgeois NOM ~ ACC’
b.	kávé	kávét	‘coffee NOM ~ ACC’
	lé	lét	‘juice NOM ~ ACC’
	izé	izét	‘thingummy NOM ~ ACC’

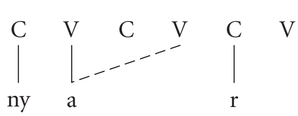
With *á*, we only find the two examples given in (21a), but with *é* there are about eighty such stems. I leave the representation of these vowels for further research.

3.2 Shortening stems

Now let me turn to the phenomenon of Shortening stems, illustrated in (15). If the representation of *á* and *é* was like that of long vowels in general, we would get the alternation given in (22a–b).

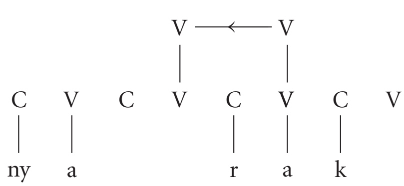
(22)

a.



nyár

b.

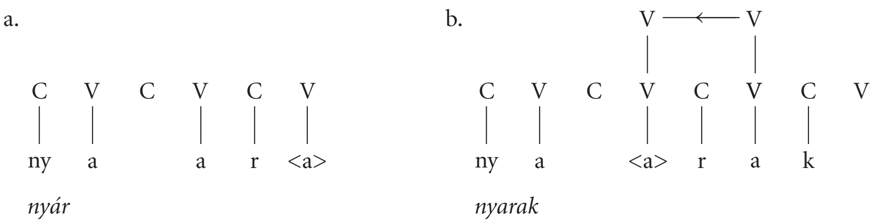


nyarak

As we have seen on the example of *hajók* in (19a), representations like (22a) have to be well-formed in Hungarian. However, there is no reason for shortening in forms like (22b) either, since the site of spreading is licensed here as well. On the basis of these representations we would thus expect no alternation (i.e. the surface forms *nyár* and **nyárak*).[9]

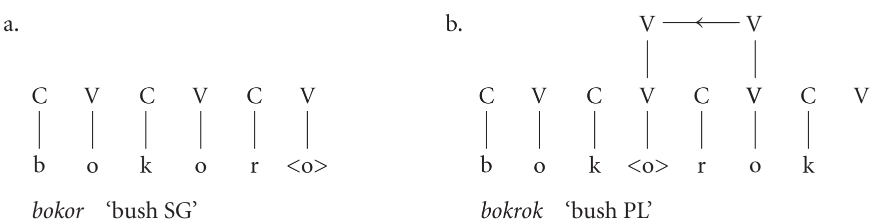
If, on the other hand, the representation of *á* and *é* is as proposed in this paper, then these forms can be analysed as given in (23). (The final vowel in (23a) is underparsed for the same reason as the one in (18b).)

(23) shortening stems



Under this analysis, shortening stems show vowel ~ zero alternation, resulting from proper government by a following full vowel. The alternation in (23) parallels the alternation shown by so-called "epenthetic" stems, given in (24), which is usually regarded as a case of vowel ~ zero alternation under proper government (cf. Törkenczy 1992).

(24) "epenthetic" stems



Note that this representation of vowel ~ zero alternation is different from the one generally proposed in Government Phonology. The usual representation was illustrated in (8), where the alternating vowel is regarded as underlyingly empty. Space limitations prevent me from giving detailed motivation for this difference (which I have done in Polgárdi 2000). However, the examples in (25) demonstrate that representing the alternating vowels as underlyingly empty is problematic in a strict CV approach.

(25)

	a. epenthetic V	b. stable V	c. stable cluster	
standard GP:	– CøCø #	– CVCø #	– CCø #	
	torony	szurony	szörny	‘sg’
	~ tornyok	~ szuronyok	~ szörnyek	‘pl’
	‘tower’	‘bayonet’	‘monster’	
strict CV:	– CøCø #	– CVCø #	– CøCø #	

The problem shown in (25) concerns the fact that there is a three-way distinction in Hungarian between word-final clusters that are sometimes broken up by a vowel, as in *torony* ~ *tornyok* in (25a), clusters that are always broken up by a vowel, as in *szurony* ~ *szuronyok* in (25b), and clusters that are never broken up by a vowel, as in *szörny* ~ *szörnyek* in (25c) – where the clusters themselves contain the same consonants (cf. Törkenczy & Siptár 1999). In a standard GP approach, these cases have separate representations. However, in a strict CV analysis, (25a) and (25c) cannot be distinguished from each other, if the alternating vowel is underlyingly empty.

Once the alternating vowel is represented as underlyingly full, on the other hand, it will have to be marked diacritically as being deletable under appropriate circumstances (i.e. if properly governed). This is inelegant. However, it explains the fact that in some languages (such as Polish) more than one vowel can alternate with zero. If these vowels are underlyingly empty, the difference between the two "epenthetic" vowels is difficult to account for.[10]

There is one further property of shortening stems, given in (26a), that is of interest here; namely, that all of these stems are also lowering (that is, they belong to the group illustrated in (18b)).

(26) a. Shortening stems (=(23))
- lowering (i.e. ending in *-a/-e*)
- alternating vowel: *á*, *é*

b. Epenthetic stems (=(24))
- non-lowering (i.e. ending in *-o/-ö/-e*)
- alternating vowel: *o, ö, e*

In addition, as I mentioned above, the alternating vowel in these cases is (almost) always either *á* or *é*. Making a survey among epenthetic stems, such as *bokor* in (24), showed that the overwhelming majority of these stems are non-lowering. (There are only about ten exceptions, apart from the nouns formed by the suffix *-alom/-elem*, which is both epenthetic and lowering.) In addition, the alternating vowel in these cases is (almost) always a short mid *o*, *ö* or *e*. This is summarised in (26b).

If we look at (23b) and (24b) again, it seems thus that there is an identity requirement on proper government in Hungarian: *a* can only be governed by *a*, while *o* can only be governed by *o*. (For a similar restriction on syncope of [a] in Palestinian Arabic, see Yoshida 1993.) However, this statement needs to be made more precise, because in its present formulation it only holds true for forms containing a consonant-initial (synthetic) suffix (such as *nya<a>ra-k* and *bok<o>ro-k*). But in fact vowel-initial synthetic suffixes (e.g. the 3pl poss *-uk*) can also properly govern the relevant *a/o* of shortening/epenthetic stems, regardless of the quality of the suffix-initial vowel (giving *nya<a>r-uk* and *bok<o>r-uk*). The restriction is therefore rather that two successive vowels can only alternate with zero, if they are identical to each other.

This statement, however, cannot be reversed. That is, the existence of two identical vowels in a row does not imply that both of them alternate with zero. This means that there are stems which look like (23) and (24), and yet they do not undergo underparsing. Some examples are given in (27).

(27) No underparsing

a. gyár gyárak 'factory sg ~ pl'
méz mézek 'honey sg ~ pl'

b. motor motorok 'engine sg ~ pl'
csömör csömörök 'disgust sg ~ pl'
zsiger zsigerek 'gut sg ~ pl'

Shortening stems and epenthetic stems thus need to be marked as such in the lexicon.

At this point, the question might arise why underparsing of *a* only occurs in examples with long vowels as in (23), and not in examples with short vowels as in (24). Checking stems that end in a $-VC_1aC_2$ # sequence, where the consonants involved are those preferred in epenthetic stems, I found only three lowering stems – which constitute the relevant context for underparsing. These are given in (28).

(28) $-VC_1aC_2$ # and lowering
vonal vonalak 'line sg ~ pl'
kazal kazlak 'rick sg ~ pl'
ajak ajkak 'lip sg ~ pl'

Interestingly, two of these (*kazal* and *ajak*) are also epenthetic – as predicted. Long *ó* and *ő*, on the other hand, are not expected to take part in this alternation, because that would involve representations as given in (22), which do not trigger shortening, as explained above.

I hope to have shown that the three peculiar properties of long *á* and *é* in Hungarian can be derived from their special representation as a sequence of two short vowels. As a consequence, they can safely occur preceding consonant clusters, and they will automatically undergo underparsing processes which apply to short non-high vowels.

4. Summary

In summary, in this paper I have argued that Hungarian syllable structure cannot be analysed using the tools of standard Government Phonology. Instead, a strict CV approach, following Lowenstamm (1996), should be taken. One argument for this comes from the fact that there are very few restrictions on what consonants can co-occur in a (word-internal) cluster. It seems thus that there is no interaction between such consonants, that is, they form a bogus cluster. Another argument is provided by the behaviour of long vowels. Apart from *á* and *é*, these cannot occur preceding consonant clusters, irrespective of the quality of the consonants involved. To be able to express this generalisation in a uniform way, all clusters need to be analysed as spurious. The prohibition on long vowels preceding a consonant cluster, in turn, follows from the principle that the target of spreading must be licensed.

In the second part of the paper, I turned to the problem of long *á* and *é*, which do occur preceding consonant clusters. These vowels are also special in that they take part in length alternations in which the other long vowels of the system do not. I have argued that these three seemingly independent facts can be derived from a single peculiarity of these vowels, namely that they are represented as a sequence of two short vowels (instead of the usual representation, where the second – empty – position is filled in by spreading from the first position).

I analysed the first phenomenon, Low vowel lengthening, as a case of domain-final shortening, which now becomes another manifestation of domain-final underparsing of non-high short vowels, argued for in Rebrus and Polgárdi (1997). The second phenomenon, that of Shortening stems, was analysed in a parallel way to the so-called epenthetic stems, that is, as underparsing resulting from proper government. I argued that vowels alternating with zero are not underlyingly empty. Instead, they are marked in the lexicon as being deletable under specific circumstances. In addition, I identified a restriction on proper government in Hungarian, namely, if two successive vowels alternate with zero, then they must be identical to each other.

On the whole thus, this analysis is rather abstract: it posits empty nuclei that cannot be heard on the surface, and sequences of short vowels that sound like ordinary long vowels, when pronounced. However, I hope to have shown that important generalisations remain unexplained, if more near-to-surface representations are employed.

Notes

* Previous versions of this paper were presented at the 'Non-existent' GP Workshop in Leiden, 1997, at the HILP4 Conference in Leiden, 1999, and at a meeting of the Linguistic Society of Szeged, 1999. I would like to thank the participants of these meetings for their helpful comments and discussion. In addition, I would like to thank Harry van der Hulst, Grażyna Rowicka, Péter Szigetvári and especially Péter Rebrus for their detailed criticism of an earlier version of this paper. All remaining errors are my own. Research for this paper was supported by the Bolyai János Scholarship of the Hungarian Academy of Sciences and a travel grant from HIL.

1. Spelling conventions: acute accents on vowels denote length (*á, é, í, ó, ú; ő, ű*)

Spelling	IPA	Spelling	IPA
ty	*c*	*sz*	*s*
gy	*ɟ*	*s*	*ʃ*
ny	*ɲ*	*zs*	*ʒ*
ly=j	*j*	*cs*	*tʃ*
		c	*ts*

2. Note that this statement is only true of monomorphemic forms. If a suffix starting with a cluster is added to a stem, the stem-final long vowel does not shorten, even if the suffix otherwise behaves as synthetic (i.e. as forming one unanalysable phonological domain with the stem): e.g. *hajó* ~ *hajónként* ~ *hajóstul* 'ship NOM ~ DISTR ~ ASSOC'. The generalisation thus only applies to the most basic lexicon of Hungarian.

3. Thanks to Péter Rebrus for drawing my attention to the above two arguments.

4. For the exact representation of the 3SG POSS ending, see Rebrus (2000: 920–931).

5. There are two suffixes that trigger *v*-Augmentation and Vowel ~ zero alternation, but do not trigger shortening: the superessive *-on* and the adjective-forming *-i*.

6. That is, I disregard the *ló* ~ *lovon* 'horse NOM ~ SUPERESS' and *ajtó* ~ *ajtaja* 'door NOM ~ 3SG POSS' type of alternations, since these involve very few lexical items. I do not deal with the *periódus* ~ *periodikus* 'period ~ periodical' type of alternations either which mostly apply preceding (derivational) suffixes of foreign origin and which often exhibit idiosyncratic properties.

7. Note that the vowel *e* occurs as both mid and low in this classification. This is so, because in the standard dialect a previous distinction between [e] and [ɛ] has merged into [ɛ] pho-

netically. However, the differences in behaviour have been preserved, [ɛ] thus manifests a dual behaviour.
Apart from the quality of the epenthetic vowel, another distinguishing property of lowering stems is that they always take an epenthetic vowel in front of the accusative suffix, even when normal stems do not require that. In this way, we can identify *méz* 'honey' as a lowering stem (surfacing as *mézet* in the accusative), while *géz* as a normal stem (giving *gézt* in the accusative), even though they form their plural in the same way.

8. Ritter (1995) and Rebrus (2000) also analyse (19b) as a case of shortening in domain-final position. However, their representation of stem-final alternating *á* and *é* parallels the representation of non-alternating *ó* and *ő*. Therefore they need to stipulate the difference between domain-final long mid and low vowels separately.

9. Ritter (1995) proposes the representations in (22) for this alternation (with the difference that the second V position is filled with the 'cold vowel'). However, her representations are not strictly CV, since only long vowels are represented by a CVCV structure, whereas consonant clusters have their usual representations as given in (1). Therefore, it is difficult to evaluate her proposal either from the perspective of standard GP, or from that of a strict CV approach. Moreover, she posits two different representations for long vowels, depending on whether they show length alternation or not.

10. On the basis of forms like *hajók* in (19a) above, we have seen that a final empty nucleus in Hungarian is strong enough to license spreading into a preceding V-position. Considering the pairs of forms in (23) and (24), however, suggests that final empty nuclei do not have the exact same licensing potential as full vowels do, since the former cannot cause underparsing (via proper government), while the latter can. Lowenstamm's (1996) proposal thus needs to be qualified for Hungarian. I leave this question open for further research.

References

Kaye, J. (1990). 'Coda' Licensing. *Phonology, 7*, 301–330.

Kaye, J. (1995). Derivations and Interfaces. In J. Durand & F. Katamba (Eds.), *Frontiers in Phonology. Atoms, Structures, Derivations* (pp. 289–332). London: Longman.

Kaye, J., Lowenstamm, J., & Vergnaud, J.-R. (1990). Constituent Structure and Government in Phonology. *Phonology, 7*, 193–231.

Kiparsky, P. (1985). Some consequences of Lexical Phonology. *Phonology, 2*, 85–138.

Lowenstamm, J. (1996). CV as the only syllable type. In J. Durand & B. Laks (Eds.), *Current Trends in Phonology: Models and Methods* (pp. 419–441). CNRS, ESRI, Paris X.

Nádasdy, Á., & Siptár, P. (1994). A magánhangzók [Vowels]. In F. Kiefer (Ed.), *Strukturális magyar nyelvtan, 2. kötet: Fonológia* (pp. 42–182). Budapest: Akadémiai Kiadó.

Papp, F. (Ed.). (1969). *A magyar nyelv szóvégmutató szótára* [Reverse-alphabetised Dictionary of the Hungarian Language]. Budapest: Akadémiai Kiadó.

Polgárdi, K. (2000). A magánhangzó ∼ zéró váltakozás a magyarban és a "csak CV"-elmélet [Vowel ∼ zero alternation in Hungarian and the "strict CV" approach]. In L. Büky & M. Maleczki (Eds.), *A mai magyar nyelv leírásának újabb módszerei IV* (pp. 229–243). Szeged: SZTE.

Rebrus, P. (2000). Morfofonológiai jelenségek [Morphophonological phenomena]. In F. Kiefer (Ed.), *Strukturális magyar nyelvtan, 3. kötet: Morfológia* (pp. 763–947). Budapest: Akadémiai Kiadó.

Rebrus, P., & Polgárdi, K. (1997). Two default vowels in Hungarian? In G. Booij & J. van de Weijer (Eds.), *Phonology in Progress – Progress in Phonology* (pp. 257–275). HIL Phonology Papers III. The Hague: Holland Academic Graphics.

Ritter, N. A. (1995). The Role of Universal Grammar in Phonology: A Government Phonology Approach to Hungarian. PhD dissertation. New York University.

Törkenczy, M. (1992). Vowel-zero alternations in Hungarian: a government approach. In I. Kenesei & Cs. Pléh (Eds.), *Approaches to Hungarian IV: The Structure of Hungarian* (pp. 157–177). Szeged: JATE.

Törkenczy, M. (1994). A szótag [The Syllable]. In F. Kiefer (Ed.), *Strukturális magyar nyelvtan, 2. kötet: Fonológia* (pp. 272–392). Budapest: Akadémiai Kiadó.

Törkenczy, M., & Siptár, P. (1999). Hungarian syllable structure: Arguments for/against complex constituents. In H. van der Hulst & N. A. Ritter (Eds.), *The Syllable: Views and Facts* (pp. 249–284). Berlin: Mouton de Gruyter.

Yoshida, S. (1993). Licensing of empty nuclei: The case of Palestinian vowel harmony. *The Linguistic Review, 10*, 127–159.

Syllable structure at different levels in the speech production process

Evidence from aphasia

Dirk-Bart den Ouden and Roelien Bastiaanse
University of Groningen, The Netherlands

Introduction

In the debate between functional and mentalist theories of language, phonology is a particularly interesting language module, for its definite relation to phonetics, the physical side of language. This relation is stronger than for any other aspect of language we have chosen to distinguish, for example syntax or semantics. When we look at phonology from this perspective, this gives rise to the question of what distinguishes phonology from phonetics. The specific question underlying this study is: "To what extent does syllable structure have a phonetic (articulatory) basis?" Adult language deficits can prove to be a good testing ground for this problem, in so far as aphasic patients can be argued to suffer from either a phonological, or a phonetic deficit.

We will present and discuss data of syllable simplifications by different types of aphasic patients. These speech errors are analyzed in terms of a syllabic template, for descriptive reasons. Before we turn to the data, however, Sections 1, 2 and 3 provide discussions of the background this study derives from.

1. Markedness in generative phonology

In the search for universal characteristics of human language, particularly boosted since the Chomskyan revolution of the 1950s and 60s, there has been an ongoing interest in the classification and recognition of linguistic structures that are more or less 'marked' than others (e.g. Chomsky & Halle 1968). What is looked at in studies devoted to the classification of markedness relations, is the relative frequency of certain structures in the languages of the world and in child language. There exist, however, many different views on what the relative markedness of certain structures over others is actually caused by, or what it means. Because of this, the term 'markedness' has become quite controversial (Battistella 1990).

Strong believers in an underlying, more abstract representation of spoken language work in a research programme aimed at giving adequate descriptions of this deeper structure. Language in its surface form is considered the window through which underlying linguistic structure can be studied. The basic frame of this underlying structure is maximally simple and universal. It is unmarked. In a simplified version of this view, the child is able to acquire language as fast as it does by starting from the simplest, already present, representation and putting more and more parameters in the complex setting, upon finding complex structures in the input.[1] Languages will have a tendency to be as simple as possible, while still being as fully communicative and subtly expressive as they can be. This conflict has the effect that, although languages differ in their surface representations, the simplest, or least marked structures will be found in most languages of the world, as basic structures will be shared to a greater extent than structures derived but further removed from the universal basis. In fact, in the case of phonology, what is devised most of the time is a markedness hierarchy of structures, in the sense that if a certain complex structure is used in a specific language, all the simpler structures within the same domain will also be present.

The best known example of this is that all languages have been claimed to have consonant-vowel (CV-)syllables, but no languages have only VC-syllables, onsets being less marked than codas (Blevins 1995).[2] Hierarchies of markedness can be considered as scales: a specific structure is not either marked or unmarked, but more marked or less marked, relative to another option. Of course, when only two structures are compared, one can speak of the marked and the unmarked candidate.

It is often suggested that children start their language acquisition (probably after a period of mere repetition) with the most basic, least marked structures (Macken 1995). A well-known proponent of this view was Roman Jakobson, who claimed that the first syllable most children acquire will be /pa/: mouth closed, mouth open.[3] This syllable, then, is regarded as the least marked syllable. Along with this example, Jakobson (1971) goes on to give a markedness hierarchy, apparently based on children's utterances, but probably also influenced by his view on segments, syllables and phonological structure. It has formed the basis for many linguistic studies.

Jakobson (1971) also claimed that "aphasic regression has proved to be a mirror of the child's acquisition of speech sounds". His idea was that the complex structure that human language is, will systematically become less complex when it is damaged, as what is learned last is first to disappear. The more the system is damaged, the more it will resemble the language of a child, using only the least marked forms. This viewpoint has formed the starting point for many research projects, but has proved to be too simple. There are many types of language deficits and the disintegration of language is more complex than was previously suggested. For example, one major difference between children acquiring language and aphasics is that the latter suffer from a deficit in a previously fully operative linguistic system, while the former are systematically building up that system (or getting to know it). The aphasic deficit, or rather the lesion site responsible for the deficit, is not bound by the system itself, as it randomly crosses functionally modular boundaries. Also, people with aphasia are adults who have grown used to many communicative strategies that are relied on in normal language production and perception. Children still have to acquire these strategies, on top of their mother tongue grammar.

Nevertheless, the logic of the general idea behind Jakobson's statement cannot be simply dismissed, especially if we are able to argue for and distinguish specific functional modules and can study patients with deficits in specific and well-defined language functions.

1.1 A phonetic basis?

If we only look at markedness from the point of view that it is a reflection of more or less 'simple' underlying structures, we run the danger of proceeding on the circular track described among others by Liberman (1996: 58): "[W]e first determine the mark from usage and then explain usage by markedness." In this sense, markedness is nothing more than a "theoretical add-on" (Ball 1996). Naturally, any next step in the direction of relevance should be an attempt to give an external explanation for why certain structures are apparently more 'simple' than others.

In looking for the factors that determine the markedness of certain structures, for example the CV- over the VC-syllable, phonologists often consider articulatory and/or perceptual characteristics of (combinations of) speech sounds. Note that taking this step is not uncontroversial. Many linguists and phonologists will defend the point that markedness is indeed a reflection of abstract preferences and basic structure, and that these are not necessarily, or not at all, related to articulation, perception, or usage. Indeed, perhaps it is the other way around and the abstract preferences have influenced the relative ease with which we articulate or perceive certain types of utterance! Points of view on this matter will mark essential differences between on the one hand, among others, Natural Phonologists, Gestural Phonologists and Functional Phonologists, and on the other hand more 'cognitive' phonologists (see e.g. Anderson 1981). Of course, most linguists are not easily classified in such general terms.

One of the problems raised by the matter of concrete versus abstract explanations is what we will call the Redundancy Argument. Once the markedness apparent from the data can be shown to be influenced entirely by articulatory or perceptual factors, let us say, the more physical side of language, or 'phonetics', the idea of abstract phonological structure is strongly undermined. This has in fact been the direction in which some phonologists have proceeded lately. Boersma (1998), for example, shows how certain aspects of phonology can be acquired without the already present basic framework proposed by traditional universalists.

However, the discussion on the physical basis of phonological constraints should not be confused with the present day chasm between behaviourism and nativism, at least not necessarily (cf. Budwig 1995). It is possible to be a functional nativist. A highly underspecified Universal Grammar, sprung from an innate drive to communicate and based on general cognitive possibilities for structure, may have been moulded by physical and functional constraints, which have then become internalized to be more effective (Bybee 1994) – as effective as they are, for example, in child language acquisition. According to this view, phonetics can be phonologized. This is the 'way out' for nativists who are confronted with more and more functional explanations for linguistic phenomena. Upon confrontation with those phenomena for which no functional basis can be found, or presented convincingly, it is the way out for functionalists.

1.2 Different levels

The introduction of multilinear phonology (e.g. Goldsmith 1979) has opened the door to the idea of markedness relations at different levels of representation. According to this view, certain phonological features, or feature settings, are less marked than others, but also combinations of certain features, segments, are less marked than other combinations. Climbing even further up the phonological hierarchy, segments may be less marked in certain syllabic positions, but more marked in others, and certain syllable structures may be marked word-internally, but unmarked at the edges of words or phrases, or in stressed positions. This yields multiple markedness hierarchies, which may sometimes come into direct conflict with each other.

Hierarchies at different phonological levels will not be constituted entirely similarly. In most theories, for example, features are either on/off (binary feature theory, e.g. Chomsky & Halle 1968) or present/not present (unary feature theory, e.g. Harris 1994). This yields either a marked situation or an unmarked situation for features. As also pointed out by Romani and Calabrese (1998), it is different for segments, which consist of bundles of features or feature settings, marked or unmarked. The markedness of segments, therefore, is more relative than that of features. In the rest of this paper, the focus will be on syllable structure.

2. Syllable structure

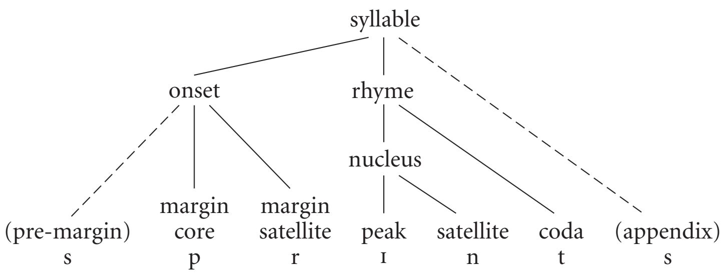
The fact that we discuss syllable *structure* here, has to do with what we think syllables are and how we think they are formed. Syllables are the building blocks of *feet*, the basic and clearest rhythmical units in language, but they are mostly regarded as the "structural units providing melodic organization" to phonological strings (Blevins 1995: 207). The possible combinations and sequences of sounds in language are constrained by syllable boundaries. Within those boundaries, only certain combinations are possible. The possibilities and constraints may be general and universal, such that no language will have syllables starting with /rt-/, but they may also be language specific, so that syllables starting with /ml-/ are possible in some languages, but not in others.

From speech error data and psycholinguistic experiments, it has been argued that syllables are not just chunks of sounds directly connected to the words that contain them, but that they should be represented as abstract structural schemas (Stemberger 1990; Sevald et al. 1995; Meijer 1996; Wilshire & Nespoulous 1997). Sounds are first of all linked to a syllabic structure (CV structure), specifying what types of sounds are in the syllable and in what order (e.g. CV, or CVC etc.). The chunk and schematic aspects of syllables may well be complementary (Sevald et al. 1995), but there has to be a level of representation where this abstract structure plays its part.

This CV structure is not just of the linear type. Syllables contain a deeper structure into onsets, nuclei and codas. All syllables must have a nuclear position (regardless of whether this position needs to be filled or not), preferably filled with a vowel. Tautosyllabic consonants in front of the nucleus form the onset, and the consonants following the nucleus make the coda. The nucleus and everything that follows make up the rhyme, so that syllables can also be divided into onsets and rhymes. The latter division is mainly made because stress and tone assignment seem to neglect the segments that make up the onset and depend only on what comes after that.

These syllable constituents can contain more than one segment, and in those cases one element heads the other. Even in frameworks where the syllable itself does not play a part as a phonological constitutent (e.g. Harris 1994), this structure is similar. A detailed template of syllabic hierarchical structure may look like the one proposed by Van Zonneveld (1988), based on Cairns and Feinstein (1982) and Van der Hulst (1984):

(1) *A syllable template*



In this particular model, satellites can only be filled with glides, liquids or nasals (i.e. sonorant consonants), the pre-margin can only be filled with the segment /s/, and the appendix with coronal voiceless obstruents (/s/, /t/) and in very rare cases /k/ or /p/. The pre-margin and appendix positions can be considered extrasyllabic – they violate binary branching and their 'behaviour' is exceptional in other ways as well (cf. Harris 1994). Positions dependent on other positions, as the pre-margin and the margin satellite are dependent on the margin core, are only filled if the position they are dependent on is filled. In this model, everything depends on the peak. In terms of markedness, the dependent positions are marked, compared to the positions they are dependent on.

The template as shown in (1) is chosen here solely for its practical advantage, because all specific syllabic positions have names one can refer to. Therefore, it works as a practical aid in the description of the data presented below. It also serves to easily show the various dependency relations present within syllables.

2.1 Cluster Reduction

If a language has consonant clusters, children acquiring this language will reduce these clusters to singletons. A number of studies have shown that this is done quite systematically, especially in onset clusters (Smith 1973; Spencer 1988b; Fikkert 1994; Gilbers & Den Ouden 1994; Paradis & Béland, to appear). The following data are samples from Gilbers and Den Ouden (1994) and from Fikkert (1994). It should be noted that they are gathered from the spontaneous speech and repetition of children and have no statistical basis, except that the linguist's ear was apparently struck by the systematicity in the utterances:

(2) a. *Child Cluster Reduction (CR), Steven* *(Gilbers & Den Ouden 1994)*

age:	target:			realization:
1;3	*bloem*	[blum]	'flower'	[bub]
1;8	*kraai*	[kraj]	'crow'	[kaj]
1;9	*klok*	[klɔk]	'clock'	[kɔk]
	stoel	[stul]	'chair'	[tuw]
	broem	[brum]	onomatopoeia	[bum]
	brrroem	[brːum]	onomatopoeia	[bumːː]
1;11	*trap*	[trap]	'stairs'	[tap]
	twee	[tʋe]	'two'	[te]
2;0	*schaap*	[sχap]	'sheep'	[bɛ] [χap]
2;2	*gloria*	[χlorija]	'gloria'	[xotija]

b. *CR-data Leonie and Jarmo* *(Fikkert 1994: 87)*

Leonie (1;9)	*slapen*	[slapən]	'to sleep'	[lapə]
Jarmo (2;3)	*slapen*	[slapən]	'to sleep'	[lapə]
Jarmo (2;3)	*slapen*	[slapən]	'to sleep'	[sapə]

We see that in CC onset clusters where the second consonant is more sonorant than the first, the second consonant is systematically deleted (3a). In sC clusters, where the second consonant is an obstruent and thus less sonorant than the fricative /s/, the /s/ is deleted (3b). Interestingly, in sC clusters where the second consonant is a liquid, the difference in sonority is apparently not sufficiently crucial, as children show more variation with respect to the reduction of these clusters (2b; 3c, d).

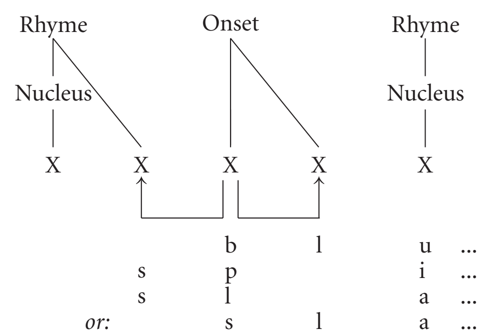
(3)

a.	C[–son]C[+son]	→ C[–son]	/trap/ →	[tap]
b.	sC[–son]	→ C[–son]	/stap/ →	[tap]
c.	sC[+son]	→ s	/slap/ →	[sap]
d.	sC[+son]	→ C[+son]	/slap/ →	[lap]

These data can be accounted for in various syllable theories, which are of course often *based* on data such as these. In the template in (1), it is obvious that the margin core position is always retained. sC(liquid) clusters can be represented with /s/ in the premargin and the liquid in the margin core, but also with /s/ in the margin core and the liquid in margin satellite position. This is apparently still optional for the struggling child, hence the error pattern in (2b).

In Government Phonology, with an empty nucleus approach (Kaye et al. 1990; Harris 1994), the representation would be as in (4):

(4) *Child Cluster Reduction in Government Phonology*



Arrows denote licensing relations. Similarly to the relative positional strength in template (1), the licensor will be retained, and the licensee will be deleted, if the cluster proves to be too difficult for the child.

An Optimality Theoretic account (Prince & Smolensky 1993) of the standard cluster reduction data may look like the tableau presented in (5), where Onset Harmony is a constraint, or a family of constraints, that formalizes the preference for syllable onsets to be as nonsonorous as possible.

(5) *Child Cluster Reduction in Optimality Theory*

klok	No-Complex	Onset	Ons-Harmony	Faithfulness
klok	* !		kl	
☞ kok			k	*
lok			l !	*
ok		* !		**

The three markedness constraints dominate Faithfulness. Again, in the case of sC(liquid) clusters, the difference in sonority between /s/ and the liquid might not be sufficiently large for Onset Harmony to be decisive at this stage of constraint (re)ranking by the child.

The above is just to illustrate that the data described and analyzed in sections to come, and especially the (syllabic) level at which they are analyzed, are independent of any specific phonological representation of syllable structure, as long as it incorporates some type of dependency or makes reference to sonority values. The point is that children apparently systematically reduce clusters to singletons, showing a preference to preserve certain prosodic elements over others. The sC(liquid) data show that it is not so much the segmental value of these elements, but their prosodic value, i.e. the syllabic position they occupy, that saves them from deletion.

3. Aphasia

Turning back to the questions posed in section 0, the phonemic paraphasias produced by aphasic speakers may serve as a good test case for the existence of a difference between an abstract level of phonology and a more concrete level of phonetics.

Phonemic paraphasias are speech errors, from which the target word is still recognizable, but in which sounds are substituted, adapted, omitted, transposed or added. The term 'phonemic' is somewhat unfortunate, as it refers to a surface level of description, i.e. in terms of phonemes, whereas it is not at all necessarily the case that the erroneous unit is the phoneme, if indeed phonemes exist at all (Harris 1998). This erroneous unit may be a single feature (hence 'adaptations'), or a syllable node, or one CV- (or X-) slot. With this acknowledged, we will use the term anyway, for its descriptive value, opposed to, for example, verbal paraphasias, in which whole words are substituted.

Patients who produce phonemic paraphasias can roughly be divided into two categories: fluent and nonfluent. In this division, nonfluent patients (often Broca's aphasics) present with an apraxia of speech and have a problem in the motor programming of a speech plan. It is important to note that they are *not* dysarthric; their speech apparatus is not paralyzed, but the difficulty lies in the timing and coordination of articulatory movements when speaking (e.g. Blumstein 1991). Patients with an apraxia of speech are characterized by their groping articulatory behaviour in the search for the correct articulation of the sound they want to produce. The number of errors in articulation is correlated with the complexity of the motor task and, through this, with the frequency of occurence of phonological structures. (Nespoulous et al. 1984, 1987; Caplan 1987; Favreau et al. 1990; Hough et al. 1994; Blumstein 1980, 1991; Code 1998).

Fluent aphasic patients presenting with phonemic paraphasias have normal articulation, but suffer from a deficit in the appropriate selection of phonemes. In a modular model of speech production, their errors originate at a deeper, phonological level of speech planning than the errors of patients with apraxia of speech. When these patients are classified into aphasic syndromes, two major types are fluent: Wernicke's aphasics and conduction aphasics. Conduction aphasia is characterized by phonemic paraphasias and a relatively (to other modalities) severe repetition disorder, with good comprehension and error-awareness, resulting in many so-called *conduites d'approches*, repeated erroneous attempts to produce the target word, not necessarily resulting in correct production (Kohn 1984). In several studies, the types of errors produced by conduction aphasics have been shown to be less systematic than, for example, the errors of Broca's aphasics with an apraxia of speech (e.g. Nespoulous et al. 1984, 1987; Bastiaanse et al. 1994; Favreau et al. 1990). Nespoulous et al. (1984) argued that conduction aphasics do not seem to be looking for a simplification of verbal output, as they found no preference for either the creation or the destruction of consonant clusters in these patients. They are regarded as suffering from a postlexical impairment in the selection and sequencing of phonemes, i.e. a deficit in string construction after lexical access (Kohn 1992; Hough et al. 1994).

Wernicke's aphasia is characterized by impaired comprehension and semantic and/or phonemic paraphasias in fluent speech. Blumstein (1980, 1991) has found that whereas Broca's aphasics show timing deficits in Voice Onset Times, Wernicke's aphasics only show minimal impairment. They mainly have problems in lexical selection.

Analogously to the fluent/nonfluent distinction, runs the distinction between posterior and anterior patients, used, for example, by Blumstein (1991). This categorization refers to lesion sites in aphasic patients, either in the left frontal lobe (Broca's aphasia), or in the left temporal lobe (Wernicke's aphasia) or left temporoparietal area (conduction aphasia).

It must be noted that many of the observations given so far in this section are controversial in the neurolinguistic community. Classification in aphasic syndromes is widely challenged, with many arguments for the alternative, a description of symptoms in case studies (e.g. Ellis & Young 1988; Démonet 1998). Also, the quite traditional dichotomy between anterior and posterior aphasic syndromes is challenged (e.g. Crosson et al. 1988). The patients in the present study, therefore, have been selected on the basis of symptoms, rather than syndromes or lesion sites. All patients produced phonemic paraphasias. The nonfluent patients were not or only very mildly dysarthric and suffered from an apraxia of speech. The fluent patients did not suffer from an apraxia of speech.

Essentially, the assumption from which this study derives is that nonfluent patients producing phonemic paraphasias suffer from a deficit that is related to a *phonetic* modality, and that fluent patients suffer from a deficit on a more abstract, underlying, *phonological* level. Even if the latter level can again be subdivided into different levels, it is at least not directly restricted by articulatory factors.

In the domain of syllabic simplification, we will investigate (a) whether, in speech production, fluent and nonfluent patients show effects that can be related to syllabic markedness as it was layed out in section 2, and (b) whether the effects are similar in both types of patients, phonetically and phonologically impaired.

4. **Method**

4.1 Subjects

Fifteen aphasic patients participated in this study. All were native speakers of Dutch. The group of nine fluent patients consisted of two females and seven males, with a mean age of 58 (range: 38–84). The nonfluent patients were six: four female and two male, with a mean age of 62 (range: 50–79). The nonfluent patients were diagnosed by their speech therapists and both authors as having an apraxia of speech in the absence (as far as this can be established) of a 'deeper' phonological deficit. The fluent patients do not have apraxia of speech. All patients produced phonemic paraphasias in spontaneous speech.

4.2 Materials

A repetition task was constructed, consisting of 114 Dutch monosyllabic words. The syllabic structures tested in these items are listed in (6). The (binary) sonority values of consonants are indicated with [+/–son] where necessary. This indication is not given where it is redundant, as for example in CC onset clusters, where the first segment is a nonsonorant and the second a sonorant consonant in Dutch.

(6) *Syllable structure categories*

		example:		
Onset singleton structures:	**C**[**–son**]V(C)	*tak*	–	/tak/
	C[**+son**]V(C)	*lak*	–	/lak/
	sV(C)	*sik*	–	/sɪk/
Onset cluster structures:	**CC**V(C)	*trap*	–	/trap/
	sC[**–son**]V(C)	*stap*	–	/stap/
	sC[**+son**]V(C)	*smak*	–	/smak/
	sCCV(C)	*strik*	–	/strɪk/
Coda singleton structures:	(C)V**C**[**–son**]	*pak*	–	/pak/
	(C)V**C**[**+son**]	*pan*	–	/pan/
	(C)V**s**	*kas*	–	/kas/
Coda cluster structures:	(C)V**C**[**+son**]**C**	*park*	–	/park/
	(C)V**C**[**+son**]**s**	*hars*	–	/hars/
	(C)V**C**[**–son**]**s**	*fiets*	–	/fits/
	(C)VsC	*mast*	–	/mast/
	(C)VC[–son]C	*pact*	–	/pakt/
	(C)VC(C)CC	*herfst*	–	/hɛrfst/

The **bold** face structures appear both in onsets and codas, though mirrored. In comparisons between onsets and codas, only these structures have been analyzed, as it would be unrealistic to compare numbers of errors made in the coda of a word such as *pact* ('pact': CVC[–son]C) with numbers of errors made in the onsets of words such as *trap* ('stairs': CC[+son]VC). In comparisons *within* constituents (onsets and codas) all structures have been analyzed.

Although there are advantages to the use of nonwords over real words in this type of experiment for reasons of item control, a very strong disadvantage is that some patients have particular problems related to nonwords. For conduction aphasics, it is sometimes quite impossible to repeat a word without being able to process it via a 'lexical route'. This may result in neologisms or silence, both of which are of no use to this study.

4.3 Procedure

Items were read from an ordered list, by one of the two authors. During the repetition task, the mouth of the researcher was shielded with the item list, so that subjects were not aided by the visible articulatory movements of the person in front of them, reading the stimuli.

4.4 Scoring and analysis

Scoring was done per syllable position, where positions were based on the template in (1). Basing the positions on this template in the data matrix was mainly done for descriptive and thus practical purposes. The results as categorized would have been no different had we used, for example, the prosodic structure proposed by Harris (1994), in the framework of Government Phonology. It yields similar predictions in prosodic simplification. Indeed, both representations are of course based on the same general observations in natural language.

Neologisms, semantic paraphasias and "no response" responses (32 in total) were not taken into account, resulting in a total valid stimulus number of 1678. Also, the errors on vowels (86 errors in 1678 utterances, or 5.1%) were not taken into consideration in this study. Only errors on consonants were scored. Many of the possible deletions in our target stimuli would result in existing words, and technically such errors would be verbal paraphasias. However, we have chosen to treat these errors as phonemic paraphasias, accidentally resembling verbal paraphasias. They were not treated any different from other valid responses.

In this study, only deletions were considered, as they are clearly a simplification of the output in absolute terms.[4] Note that it is essentially not the question here whether and which aphasic patients simplify syllable structure, but rather which syllabic positions are most vulnerable in the case of simplification.

Only the first attempts of patients were considered. Wilcoxon tests were conducted on the proportions of deletions in different syllable positions present in target items. Proportions were derived from the total number of occurences of a position in the original targets, so that additions did not raise the total number, and thus lower the proportion of deletions.

5. Results

In the results presented below, significance is established at $p < 0.05$. Probability values follow from two-tailed Wilcoxon tests on the individual scores of patients. For the sake of readability of this article, we have chosen to present only the values of significant differences in the text, while all other z- and p-values are listed in Appendix A.

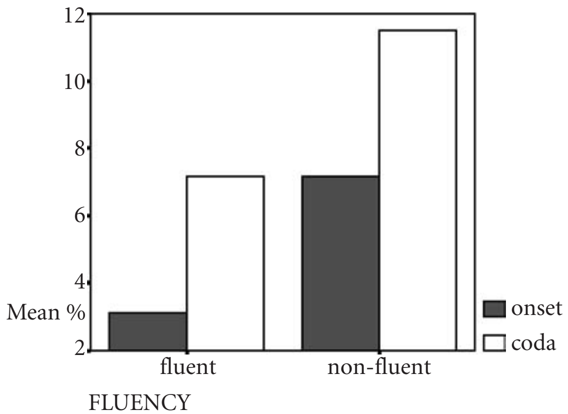
All graphs show the average percentages of deletions in a certain position, relative to the number of times the position is filled in the target word. In the graphs, PREMAR means 'premargin'; MARCOR means 'margin core';

MARSAT means 'margin satellite'; and CODSAT means the 'sonorant coda position'.

5.1 Onsets vs. codas

Figure (7) shows the average percentages of deletions in onsets and codas of fluent and nonfluent patients. Both groups of patients make more errors in codas than in onsets.

(7) *Onsets vs. codas per group*

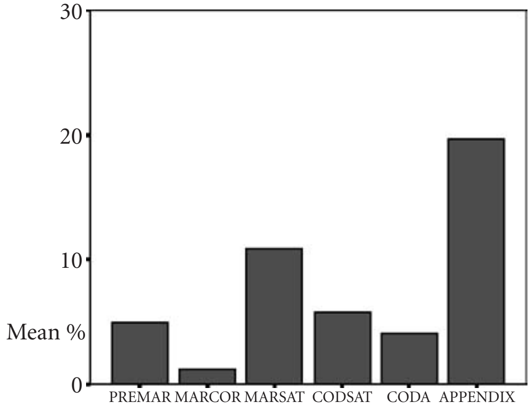


For fluent patients, this difference is significant (z = –2.240; p = 0.025), but for nonfluent patients it is not (z = –0.524; p = 0.6).

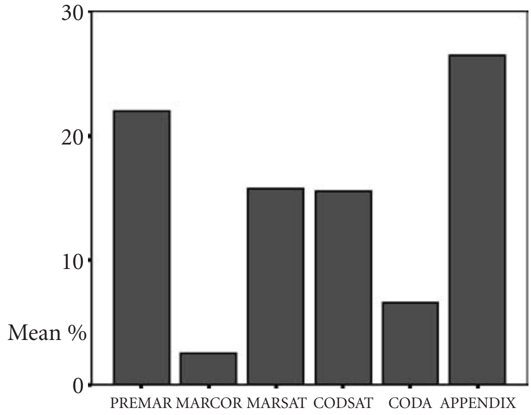
5.2 Positional effects within clusters

Figures (8) and (9) show the average percentages of deletions per position of the fluent patients and the nonfluent patients.

(8) *Fluent patients' deletions by position*



(9) *Nonfluent patients' deletions by position*



For the fluent patients, the following effects are significant:

I. more deletions of the margin satellite than of the margin core (z = –2.197; p = 0.028)
II. more deletions of the appendix than of the coda position (z = –2.666; p = 0.008)
III. more deletions of the appendix than of the sonorant coda position (z = –2.310; p = 0.021)

For the nonfluent patients, significance is reached for the following effects:

I. more deletions of the margin satellite than of the margin core (z = –2.201; p = 0.028)

II. more deletions of the appendix than of the coda position (z = –2.201; p = 0.028)

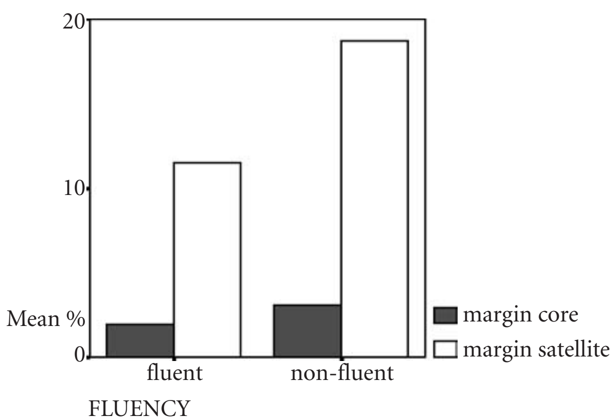
The great average percentage of deletion errors in the premargin of nonfluent patients is mainly caused by one patient in that group, who produced a deletion error in 94.4% of the time this position occurred in the target word. As the Wilcoxon test takes into account the individual scores of the subjects, the difference between the premargin and the margin core, which is quite striking in the graph, proves not to be significant for the nonfluent patients. The difference between the sonorant coda position (CODSAT) and the 'normal' coda position does not reach significance for either group. Particularly the nonfluent patients, however, have a tendency to delete the coda satellite position, for example filled with /n/ in '*print*'.

5.2.1 *Consonant clusters in onsets*

In order to investigate the relative vulnerability to deletion of different syllabic positions in more detail, we will now focus on specific consonant clusters. This also provides us with a chance to analyze the data with even less dependence on the syllabic representation on which we have chosen to model our target items. In the following graphs, the bars belonging to positions that template (1) predicts to be weak are light, and the strong positions are dark.

Figure (10) shows the average percentages of deletions of both groups of patients in CC onset clusters, of the type '***p**lot*'.

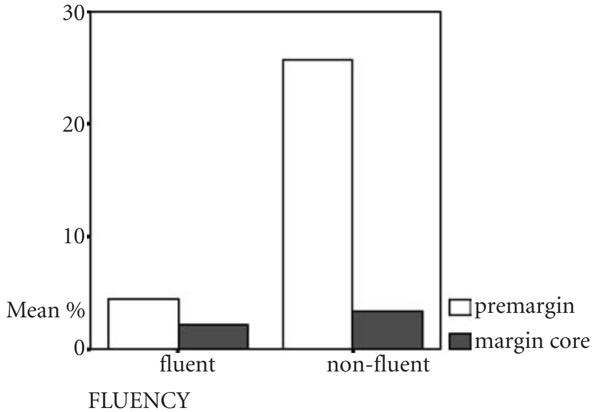
(10) *Deletions in CC onsets*



For both fluent (z = –2.032; p = 0.042) and nonfluent (z = –2.041; p = 0.041) subjects, the effects are significant: more deletions in the margin satellite.

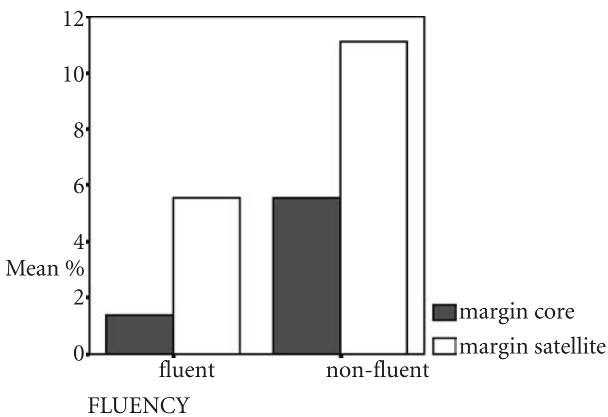
Figure (11) shows the average percentages of deletions in sC[–son] onset clusters, of the type '*stap*'. Neither of the differences is significant.

(11) *Deletions in sC[–son] onsets*



Average percentages of deletions in sC[+son] onset clusters, of the type '*slap*', are depicted in Figure (12).

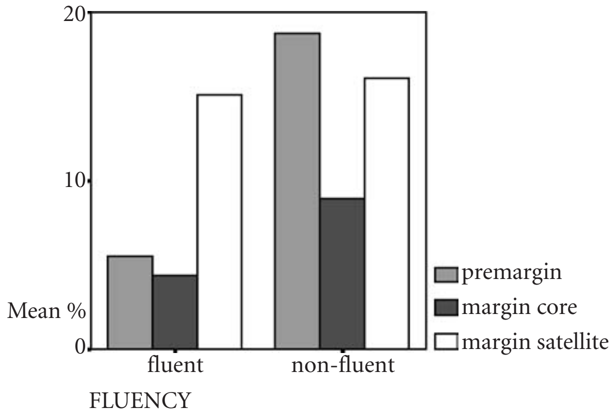
(12) *Deletions in sC[+son] onsets*



Neither of the differences is significant.

Figure (13) depicts the average percentages of deletions in sCC onset clusters, of the type '**str**af'. Again, none of the differences is significant.

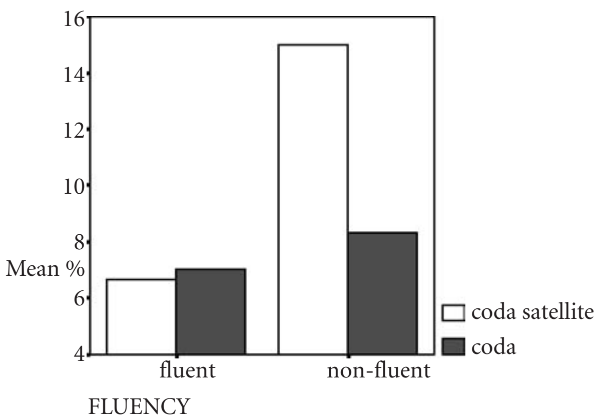
(13) *Deletions in sCC onsets*



5.2.2 *Consonant clusters in codas*

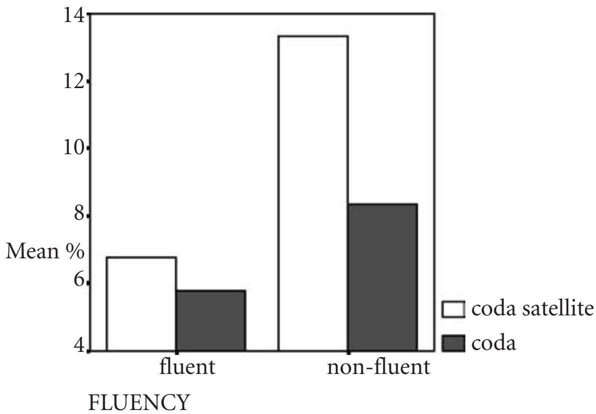
Figure (14) shows the percentages of deletions in C[+son]C codas, of the type '*park*'. For the fluent patients, the sonorant coda position and the coda position are equally vulnerable to deletion. Nonfluent patients show a tendency to delete the segment in the sonorant coda position more often than that in the coda position.

(14) *Deletions in C[+son]C codas*



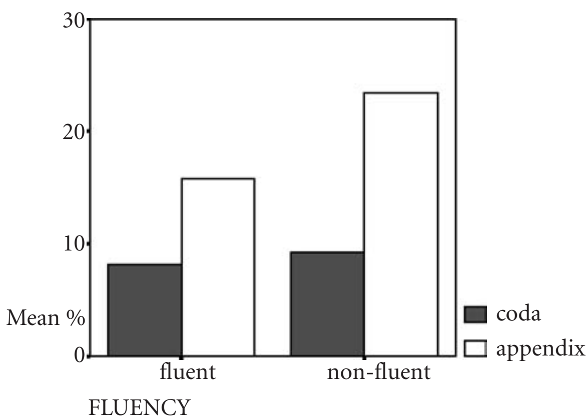
If the second consonant in this type of onset cluster is /s/, as in '*hars*', the picture is the same (Figure (15)).

(15) *Deletions in C[+son]s codas*



Coda clusters of type C[–son]s ('*muts*') are represented differently in the model, as they violate the sonority hierarchy (when regarded as one syllable). C[–son] is placed in the coda position, and /s/ in the appendix. Average percentages of deletions in these positions are shown in Figure (16).

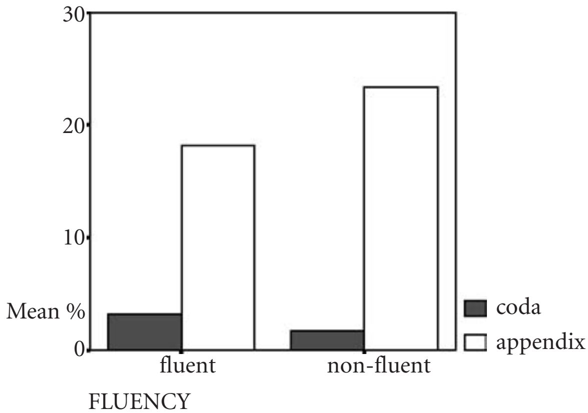
(16) *Deletions in C[–son]s codas*



Neither of the differences in Figure (16) reaches significance.

Figure (17) depicts the average percentages of deletions in sC coda clusters, of type '*mist*', where, in our model (1), /s/ is placed in the coda position and the following consonant in the appendix.

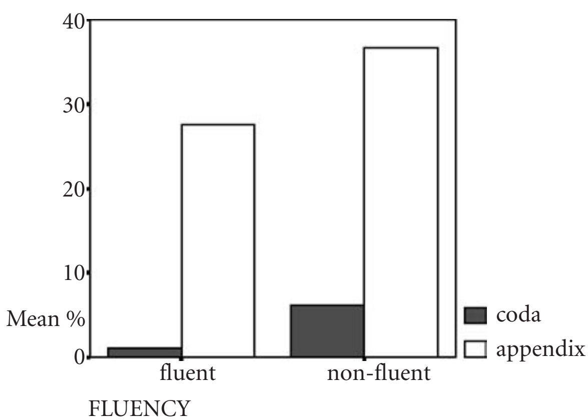
(17) *Deletions in sC codas*



Fluent patients produce significantly more deletions in the appendix position than in the coda position (z = –2.588; p = 0.01), but this is merely a tendency (z = –1.604; p = 0.109) for nonfluent patients.

C[–son]C coda clusters, of type '*pact*', which have the same representation, evoke similar effects, depicted in Figure (18).

(18) *Deletions in C[–son]C codas*



Again, the effect is significant for fluent patients (z = –2.527; p = 0.012), but not for nonfluent patients (z = –1.581; p = 0.114).

The category of (C)VC(C)CC codas was not regarded and analyzed as one homogeneous group, because it consists of words that vary in their representation, such as '*herfst*' and '*markt*'. The first of these two examples has two segments in the appendix, whereas the second only has one segment in that position.

6. Discussion

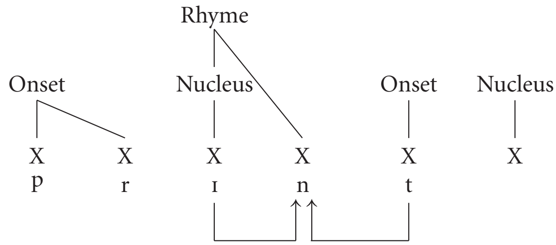
None of the significant effects or the nonsignificant tendencies found in the fluent and nonfluent aphasics' utterances is in the opposite direction as could be expected from the syllabic representation used, which is based on, among others, child cluster reduction data and phonological structures and processes gathered from languages of the world.

In some cases, the lack, or weakness of effects is noteworthy. The premargin position, occupied by /s/ in, for example, '*step*' proves not to be as vulnerable to reduction as might be expected from its representation. /s/ violates the sonority hierarchy in this position. In Government Phonology (e.g. Harris 1994) it belongs to a previous syllable, licensed by the consonant that follows it (/t/ in '*step*'). Also, when /s/ is less sonorant than the following consonant in an onset, as in '*slap*', there are only tendencies for both groups of patients to delete the second consonant, which, in our model (1), is indeed a dependant of /s/ in the margin core position. All tendencies are in the direction of less marked structures, but perhaps /s/ appears to behave with somewhat more variability than other types of consonants in the onset. In section 2.2, we saw that sC[+son] onset clusters also appear to lead to more variable behaviour in children reducing clusters.

Regarding the sC[+son] onset clusters, with /s/ in the margin core, fluent and nonfluent aphasic speakers behave quite similarly, with respect to the relative (nonsignificant) difference between margin core and margin satellite positions, with more deletions in the satellite position. With respect to sC[–son] onset clusters, with /s/ in the premargin, the effects of both groups are in the direction of more deletions in the premargin, as expected from the perspective of positional markedness, but these effects are not significant, and for the fluent group it is very small.

In codas, fluent patients make an equal amount of deletions in the sonorant coda position (the coda satellite) as in the coda position, while nonfluent patients have a tendency to delete the sonorant coda position (/n/ in '*print*') more often. Phonological theory is less straightforward regarding the relative markedness of codas, than it is for onsets. From template (1) and the Government Phonology representation in (19), we would expect /t/ in /prɪnt/ to be stronger than /n/.

(19) *Representation of 'print' according to Government Phonology*



In this representation, /n/ is licensed by /ɪ/, the nucleus, and by /t/, the following onset, which also properly governs /n/ (Harris 1994: 168). A licensee is naturally expected to be weaker than its licensor.

However, Clements (1990) claims that codas do in fact have a 'preference' to be sonorant, which makes /pɪ**n**/ less marked than /pɪ**t**/. This Sonority Dispersion Principle is in disagreement with, for example, Donegan and Stampe's (1978) Theory of Maximal Contrast (cf. Christman 1992b; Romani & Calabrese 1998). The claim is based on an analysis of the frequency of structures that are present in languages in the world. From a purely segmental perspective, most phonologists would argue that in the domain of consonants, the least sonorant consonants are the least marked and the more sonorant consonants are more marked. From a prosodic perspective, however, it may be the case that more sonorant consonants in codas are preferred over less sonorant consonants (cf. section 1.3). Possibly, this conflict in codas leads to the balance in the results of fluent patients when the sonorant positions are compared with the nonsonorant positions. In onsets, the conflict does not exist, because onsets have a preference to be as nonsonorant as possible, according to both the Sonority Dispersion Principle and the Theory of Maximal Contrast.

Apart from these two effects, or non-effects, both groups of patients behave similarly and as expected from the syllabic representation used as a background to this study, whether in tendencies or significant effects. This deals with our two questions as stated at the end of section 3. In light of the discussion about phonetics versus phonology, these results give rise to the claim that syllable structure, with positional prominence relations, does not only play a role at the articulatory (phonetic) level, but comes into play on an earlier, phonological level in the speech production process. Of course, this does not mean that the structures present at both levels have to be independent. It seems rather the case that the two levels are highly interdependent, where either articulatory constraints have influenced phonological representation, or, as theoretically possible but less likely, the other way around. To us, these results provide a hint in the direction of functional nativism as described in section 1.2.

A big 'however' says that patients with an apraxia of speech may not have an articulatorily based, or phonetic problem at all. Studies have indeed indicated before that apraxia of speech may be more of a phonological problem, where errors are highly influenced by underlying phonological representations (e.g. Dogil & Mayer 1998). Code and Ball (1988), for example, argue for a *phonetology*, with a phonological component, a cognitive phonetic component and an articulatory phonetic component. Apraxia of speech, in this theory, is an impairment at the cognitive phonetic component. Although these authors specifically argue that markedness reflects articulatory properties and not linguistic ones, the data presented here at least show that markedness effects are present both at the phonological level and at the (cognitive) phonetic level. Neurolinguists are still debating (Code 1998).

7. Conclusion

We have reported on a study into the treatment of syllable structure by phonetically and phonologically impaired aphasic patients. Regarding the relative similarity of the error types of both groups of patients it was argued that syllable structure plays a part in an early stage of the speech production process, but that it also has an articulatory basis. Syllable markedness effects cannot only be considered a direct result of articulatory factors, as the hierarchic structures responsible for the effects must also be present at a more abstract, underlying level of processing.

Notes

1. Or, alternatively, by giving greater power to constraints that allow the output to be more complex, so as to match it to the surface form of the language the child is acquiring.

2. More specifically, not having an onset filled with segmental material makes a syllable more marked, or less likely to occur than syllables in which the onset is filled.

3. From a segmental, or rather featural perspective, the first syllable should be /ta/, as the coronal place of articulation of /t/ is generally considered the least marked and therefore default candidate.

4. A simplification, here, is literally a reduction in the number of phonemes, or articulatory gestures the subject has to make. In this terminology, a simplification may lead to a more

marked structure, where for example a CVC syllable is reduced to a VC syllable, 'lacking' an onset. Similarly, addition of a segment may result in a less marked syllable structure, as, for example, with the addition of a vowel within a complex onset (/trap/ → [tərap]). However, this would not constitute a reduction in *absolute* terms.

Appendix A

Onsets versus codas (Figure (7)):
Fluent: z = –2.240; p = 0.025
Nonfluent: z = –0.524; p = 0.6

Positional effects all syllables
Fluent patients (Figure (8))
PREMAR vs. MARCORE:
z = –1.753; p = 0.080
MARCORE vs. MARSAT: z = –2.197; p = 0.028
PREMAR vs. MARSAT: z = –1.439; p = 0.15
CODSAT vs. CODA: z = –1.120; p = 0.263
CODA vs. APPENDIX: z = –2.666; p = 0.008
CODSAT vs. APPENDIX: z = –2.310; p = 0.021

Nonfluent patients (Figure (9))
PREMAR vs. MARCORE:
z = –1.572; p = 0.116
MARCORE vs. MARSAT: z = –2.201; p = 0.028
PREMAR vs. MARSAT: z = –0.524; p = 0.6
CODSAT vs. CODA: z = –1.153; p = 0.249
CODA vs. APPENDIX: z = –2.201; p = 0.028
CODSAT vs. APPENDIX: z = –1.363; p = 0.173

*CC onsets, **trap** (Figure (10))*
Fluent: z = –2.032; p = 0.042
Nonfluent: z = –2.041; p = 0.041

*sC[–son] onsets clusters, **stap** (Figure (11))*
Fluent: z = –0.816; p = 0.414
Nonfluent: z = –1.461; p = 0.144

*sC[+son] onsets **slap** (Figure (12))*
Fluent: z = –1.089; p = 0.276
Nonfluent: z = –0.816; p = 0.414

*sCC onsets, **straf** (Figure (13))*
Fluent
PREMAR vs. MARCORE:
z = –0.412; p = 0.680
MARSAT vs. MARCORE:
z = –1.511; p = 0.131
PREMAR vs. MARSAT:
z = –1.577; p = 0.115
Nonfluent
PREMAR vs. MARCORE:
z = –0.378; p = 0.705
MARSAT vs. MARCORE:
z = –0.813; p = 0.416
PREMAR vs. MARSAT:
z = –0.135; p = 0.893

*C[+son]C codas, **park** (Figure (14))*
Fluent: z = –0.378; p = 0.705
Nonfluent: z = –1.414; p = 0.157
*C[+son]s codas, **mals** (Figure (15))*
Fluent: z = –0.108; p = 0.914
Nonfluent: z = –1.134; p = 0.257

*C[–son]s codas, **muts** (Figure (16))*
Fluent: z = –0.984; p = 0.325
Nonfluent: z = –1.483; p = 0.138

*sC codas, **mist** (Figure (17))*
Fluent: z = –2.588; p = 0.01
Nonfluent: z = –1.604; p = 0.109)

*C[–son]C codas, **pact** (Figure (18))*
Fluent: z = –2.527; p = 0.012)
Nonfluent patients: z = –1.581; p = 0.114.

References

Anderson, S. R. (1981). Why phonology isn't "natural". *Linguistic Inquiry*, *12* (4), 493–539.

Ball, M. J. (1996). *Evidence from vowel disorders for the nature of minimal phonological units in acquisition.* Paper presented at the Fifth Annual Conference of the International Clinical Phonetics and Linguistics Association. München, September 16–18, 1996.

Bastiaanse, R., Gilbers, D. G., & Van der Linde, K. (1994). Sonority substitutions in Broca's and conduction aphasia. *Journal of Neurolinguistics*, *8* (4), 247–255.

Battistella, E. L. (1990). *Markedness: The Evaluative Superstructure of Language*. Albany: State University of New York Press.

Blevins, B. (1995). The syllable in phonological theory. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 206–244). Oxford: Blackwell Publishers.

Blumstein, S. E. (1991). Phonological aspects of aphasia. In M. T. Sarno (Ed.), *Acquired Aphasia*. San Diego: Academic Press.

Blumstein, S. E., Cooper, W. E., Goodglass, H., Statlender, S., & Gottlieb, J. (1980). Production deficits in aphasia: A voice-onset time analysis. *Brain and Language*, *9*, 153–170.

Boersma, P. (1998). Functional Phonology: Formalizing the Interactions between Articulatory and Perceptual Drives. PhD Dissertation, University of Amsterdam.

Budwig, N. (1995). *A Developmental-Functionalist Approach to Child Language*. Mahwah, NJ: Lawrence Erlbaum.

Bybee, J. (1994). A view of phonology from a cognitive and functional perspective. *Cognitive Linguistics, 5*, 285–305.

Cairns, C., & Feinstein, M. (1982). Markedness and the theory of syllable structure. *Linguistic Inquiry*, *13*, 193–226.

Caplan, D. (1987). *Neurolinguistic and Linguistic Aphasiology; an Introduction*. Cambridge Studies in Speech Science and Communication. Cambridge: Cambridge University Press.

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper &

Christman, S. S. (1992). Uncovering phonological regularity in neologisms: Contributions of sonority theory. *Clinical Linguistics & Phonetics*, *6* (3), 219–247.

Clements, G. N. (1990). The role of sonority in core syllabification. In J. Kingston & M. E. Beckman (Eds.), *Papers in Laboratory Phonology I. Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press.

Code, C. (1998). Models, theories and heuristics in apraxia of speech. *Clinical Linguistics and Phonetics*, *12* (1), 47–65.

Code, C., & Ball, M. J. (1988). Apraxia of speech: the case for a cognitive phonetics. In M. J. Ball (Ed.), *Theoretical Linguistics and Disordered Language* (pp. 152–167). London & Sydney: Croom Helm.

Crosson, B., Novack, T. A., & Trenerry, M. R. (1988). Subcortical language mechanisms: Window on a new frontier. In H. A. Whitaker (Ed.), *Phonological Processes and Brain Mechanisms*. New York etc.: Springer-Verlag.

Démonet, J. F. (1998). Tomographic Brain Imaging of Language functions: Prospects for a New Brain/Language Model. In B. Stemmer & H. A. Whitaker (Eds.), *Handbook of Neurolinguistics* (pp. 131–142). San Diego: Academic Press.

Dogil, G., & Mayer, J. (1998). Selective phonological impairment: a case of apraxia of speech. *Phonology*, *15* (2).

Donegan, P. J., & Stampe, D. (1979). The study of natural phonology. In D. A. Dinnsen (Ed.), *Current Approaches to Phonological Theory* (pp. 126–173). Bloomington: Indiana University Press.

Ellis, A. W., & Young, A. W. (1988). *Human Cognitive Neuropsychology*. London etc.: Lawrence Erlbaum Associates Ltd.

Favreau, Y., Nespoulous, J.-L., & Lecours, A. R. (1990). Syllable structure and lexical frequency effects in the phonemic errors of four aphasics. *Journal of Neurolinguistics*, *5* (2/3), 165–187.

Fikkert, P. (1994). On the Acquisition of Prosodic Structure. PhD dissertation, University of Leiden.

Gilbers, D. G., & Den Ouden, D. B. (1994). Compensatory lengthening and cluster reduction in first language acquisition: A comparison of different analyses. In A. de Boer, H. de Hoop, & H. de Swart (Eds.), *Language and Cognition 4, Yearbook 1994 of the research group for Theoretical and Experimental Linguistics of the University of Groningen*, 69–82.

Harris, J. (1994). *English Sound Structure*. Oxford: Blackwell.

Harris, J. (1998). Phonological universals and phonological disorders. In E. G. Visch-Brink & R. Bastiaanse (Eds.), *Linguistic Levels in Aphasia* (pp. 91–117). San Diego: Singular Publishing Group.

Hough, M. S., Demarco, S., & Farler, D. (1994). Phonemic retrieval in conduction aphasia and Broca's aphasia with apraxia of speech: underlying processes. *Journal of Neurolinguistics*, *8* (4), 235–246.

Jakobson, R. (1971). *Studies on Child Language and Aphasia*. Janua Linguarum, Series Minor, 114. The Hague, Paris: Mouton.

Kaye, J., Lowenstamm, J., & Vergnaud, J.-R. (1990). Constituent structure and government in phonology. *Phonology*, *7*, 193–232.

Kohn, S. E. (1984). The nature of the phonological disorder in conduction aphasia. *Brain and Language*, *23*, 97–115.

Kohn, S. E. (1992). Conclusions: Towards a working definition of conduction aphasia. In S. E. Kohn (Ed.), *Conduction Aphasia*. Hillsdale etc.: Lawrence Erlbaum Associates.

Liberman, A. (1996). Phonological markedness and a plea for Useful Linguistics. In E. Andrews & Y. Tobin (Eds.), *Toward a Calculus of Meaning: Studies in Markedness, Distinctive Features and Deixis* (pp. 55–69). Amsterdam: Benjamins.

Macken, M. A. (1995). Phonological acquisition. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 671–696). Oxford: Blackwell Publishers.

Meijer, P. J. A. (1996). Suprasegmental structures in phonological encoding: The CV structure. *Journal of Memory and Language*, *35*, 840–853.

Nespoulous, J.-L., Joanette, Y., Béland, R., Caplan, D., & Lecours, A. R. (1984). Phonological disturbances in aphasia: Is there a "markedness effect" in aphasic phonemic errors? In F. C. Rose (Ed.), *Advances in Aphasiology, 42, Progress in Aphasiology*. London: Raven Press.

Nespoulous, J.-L., Joanette, Y., Ska, B., Caplan, D., & Lecours, A. R. (1987). Production deficits in Broca's and conduction aphasia: Repetition vs. reading. In E. Keller & M. Gopnik (Eds.), *Motor and Sensory Processes in Language*. Hillsdale, NJ: Erlbaum.

Paradis, C. & Béland, R. (2002). Syllabic constraints and constraint conflicts in loanword adaptations, aphasic speech and children's errors. In B. Laks & J. Durand (Eds.), *Phonetics, Phonology and Cognition*. Oxford: Oxford University Press.
Prince, A., & Smolensky, P. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers University Center for Cognitive Science Technical Report *2*.
Romani, C., & Calabrese, A. (1998). Syllabic constraints in the phonological errors of an aphasic patient. *Brain & Language*, *64*, 83–121.
Sevald, C. A., Dell, G. S., & Cole, J. S. (1995). Syllable structure in speech production: Are syllables chuncks or schemas? *Journal of Memory and Language*, *34*, 807–820.
Spencer, A. (1988). A phonological theory of phonological development. In M. J. Ball (Ed.), *Theoretical Linguistics and Disordered Language* (pp. 115–151). London & Sydney: Croom Helm.
Stemberger, J. P. (1990). Wordshape errors in language production. *Cognition*, *35*, 123–157.
Van der Hulst, H. (1984). *Syllable Structure and Stress in Dutch*. Dordrecht: Foris.
Van Zonneveld, R. (1988). Two level phonology: structural stability and segmental variation in Dutch child language. In F. Van Besien (Ed.), *First Language Acquisition, ABLA Papers, 12*, 129–162.
Wilshire, C. E., & Nespoulous, J.-L. (1997). The role of syllabic frequency in aphasic phonological errors. *Brain & Language*, *60* (1), 147–150.

Metrical structure

Quantity-sensitivity of syllabic trochees revisited

The case of dialectal gemination in Finnish*

Heli Harrikari
University of Helsinki

1. Introduction

Quantity-insensitivity has been considered as one of the crucial properties of syllabic trochees[1] (Hayes 1995). Thus, unlike the other foot forms – moraic trochees and iambs – syllabic trochees group syllables into sets of two, irrespective of syllable weight.[2] This paper, however, demonstrates that in order to account for certain features of dialectal gemination in Finnish, the quantity-insensitivity of syllabic trochees must be re-evaluated.

In various eastern and northern Finnish dialects (see Rapola 1961 for the detailed lists of the geographic areas), single word-medial consonants are frequently geminated. For example, /kalaa/[3] 'fish, sg. partitive' appears as [kallaa][4] (Rapola 1961, 1966; Jalava 1971; Paunonen 1973, 1974; Holman 1976; Kangasmaa-Minn 1978; Nahkola 1987; Palander 1987). Gemination (a fortition process), which started as phonetic variation, has been historically described as a response to the appearance of complex nuclei in non-initial syllables, which derived from the deletion of an intervening consonant (as in [kál.la-a],[5] from [ká.la-a], from [ká.la.-ða] 'fish, sg. partitive'). As a consequence of this diachronic change, the metrical balance between the (main-) stressed syllable and the following unstressed syllable was disturbed, since the stressed syllable was now lighter than the following unstressed one. Gemination has thus been seen as a reconciliation of the metrically undesirable situation, as an attempt to increase the weight of the stressed syllable.

This paper provides a formal synchronic analysis for dialectal gemination. I will demonstrate – through an Optimality Theoretic (OT, Prince &

Smolensky 1993; McCarthy & Prince 1993) analysis – that earlier historically-based accounts of dialectal gemination as a mean of increasing the weight of the stressed syllable are inadequate. The analysis shows that synchronically, gemination is not a matter of a specific syllable, but that of the higher prosodic unit, the foot. I will show that gemination is, in fact, an attempt to achieve foot-internal balance, a pattern which derives from the interaction of the concept of a more optimal trochee and positional faithfulness. The analysis demonstrates that different types of syllabic trochees form an intrinsic ranking, which indicates more desirable trochees as well as less optimal ones, crucially relying on quantity-sensitivity by syllabic trochees. Furthermore, the analysis illustrates the dominant position of positional faithfulness over syllable markedness. Only through this ranking can we explain the fact that gemination appears as the optimal means of improving the foot balance, even though at the same time, gemination creates cross-linguistically marked syllable structures.

The paper is organized as follows. First, Section 2 introduces the data. In Section 3, I will propose an OT analysis for dialectal gemination by showing how gemination results from the interaction of the different types of syllabic trochees and positional faithfulness. Crucially, I will demonstrate the need for an internal preference of syllabic trochees, as well as illustrate the necessity of the crucial ranking of positional faithfulness and syllable markedness. Finally, Section 4 summarizes the analysis.

2. Data description

Dialectal gemination occurs in certain morphological environments, such as the singular and plural partitive, singular illative of nouns, and in certain verb forms, such as the first infinitive and the third person singular present tense. The markers (suffixes) of these morphological categories are vowel-initial, and, when attached to the stem, they form a long vowel or a diphthong with the stem-final vowel. Gemination is exemplified in (1) (Rapola 1966). Forms including geminates are listed alongside their stems, which illustrate non-geminated forms. The non-geminating stems are used with most morphological endings, and are thus generally considered the "elsewhere" case. (They are also used as the input form in the analysis in Section 3.) Since the marker of every morphological category in question is a change in the length of the stem-final vowel (followed by a consonant in certain cases), I characterize the morpheme as a bare mora.

(1)	Stem[6] (non-geminated)	Geminated form		Gloss
	/jaka-/	[já**k**.**k**aa]	< /jaka+μ/	's/he/it shares'
	/sano-/	[sá**n**.**n**oo]	< /sano+μ/	's/he/it says'
	/mene-/	[mé**n**.**n**ee]	< /mene+μ/	's/he/it goes'
	/pitä-/[7]	[pí**t**.**t**ää]	< /pitä+μ/	'must'
	/liha-/	[lí**h**.**h**aa]	< /liha+μ/	'meat, sg. partitive'
	/pata-/	[pá**t**.**t**aa]	< /pata+μ/	'pot, sg. partitive'
	/kala-/	[ká**l**.**l**aa]	< /kala+μ/	'fish, sg. partitive'
	/save-/	[sá**v**.**v**ee]	< /save+μ/	'mud, sg. partitive'
	/kuva-/	[kú**v**.**v**aa]	< /kuva+μ/	'picture, sg. partitive'
	/oma-/	[ó**m**.**m**aa]	< /oma+μ/	'own, sg. partitive'
	/iso-/	[í**s**.**s**oon]	< /iso+μn/	'big, sg. illative'
	/nime-/	[ní**m**.**m**een]	< /nime+μn/	'name, sg. illative'
	/kumartele-/	[kú.mar.tè**l**.**l**ee]	< /kumartele+μ/	's/he/it nods, frequentative'
	/alentako-/	[á.len.tà**k**.**k**oon]	< /alentako+μn/	'let him lower, causative'

Gemination has the following properties: the onset consonant of the unstressed syllable is geminated when preceded by a light, stressed syllable (i.e. (C)V), and the unstressed syllable itself includes a long vowel or a diphthong (i.e. a complex nucleus). Furthermore, the data in (1) illustrate that the environment in which gemination occurs is a sequence of two syllables, [CV] and [CVV].[8] This crucial observation of bisyllabicity suggests that the domain of gemination may be a foot. This hypothesis is supported by the fact that the Finnish foot type is generally considered to be a syllabic trochee (Kenstowicz 1994a; Hayes 1995; McCartney 1999). On the contrary, gemination does not take place in contexts in which an unstressed syllable would be strengthened, as shown in (2).

(2)	/omena-/	[ó.me.naa]	< /omena+μ/	'apple, sg. partitive'
		*[ó.me**n**.**n**aa]		
	/salama-/	[sá.la.maa]	< /salama+μ/	'lightning, sg. partitive'
		*[sá.la**m**.**m**aa]		

Having examined the relevant data, let us now turn to a formal Optimality Theoretic analysis of gemination.

3. Analysis

In this section, I will propose an Optimality Theoretic analysis of dialectal gemination, by which gemination is an attempt to form a more balanced syllabic trochee.[9] I will demonstrate how gemination results from the interaction of trochee-oriented constraints and positional faithfulness. The analysis suggests that syllabic trochees are, in fact, quantity-sensitive in the dialects in question, contrary to earlier assumptions of the quantity-insensitivity of syllabic trochees.

Since my analysis crucially relies on the concept of *a more optimal trochee*, I will first discuss different types of syllabic trochees in order to establish which types of trochees are preferred and which ones are disfavored; in other words, the intrinsic ranking of syllabic trochee types. (For a similar type of discussion on different foot types based on the Grouping Harmony in a pre-OT framework, see Prince 1991.)

3.1 Internal balance of syllabic trochees

It is well known that trochees prefer its syllables to be even in weight (McCarthy & Prince 1986; Prince 1991; Kager 1993; Hayes 1995). It follows that syllabic trochees of the form (CV́.CV), (CV́C.CVC), and (CV́V.CVV) are favored, while feet containing two syllables of unequal weight are less desirable. The aforementioned even trochees have an intrinsic ranking: (CV́.CV) is the most optimal trochee, due to the fact that it consists of two light syllables, while (CV́V.CVV) is the worst of the even trochees, because of the heaviness of the two syllables. Following this reasoning, the hierarchy of the even trochees, (CV́.CV) ≻ (CV́C.CVC) ≻ (CV́V.CVV), is established. (The same hierarchy expressed in OT terms, *(CV́V.CVV) >> *(CV́C.CVC) >> *(CV́.CV).)

Given the idea of superiority of certain types of syllabic trochees, three Optimality Theoretic constraint hierarchies can be established, reflecting the internal balance of syllabic trochees. The least favored trochee appears in the highest position, furthest to the left in the hierarchies in (3).

(3) a. *(CV́.CVV) >> *(CV́.CVC) >> *(CV́.CV)
  b. *(CV́V.CV) >> *(CV́V.CVC) >> *(CV́V.CVV)
  c. *(CV́C.CVV), *(CV́C.CV) >> *(CV́C.CVC)

The hierarchy in (3a) shows trochees with the first syllable light. The trochee of the shape (CV́.CVV) is the most undesirable trochee, since the weight difference between the two syllables is the largest possible. Furthermore, the trochee

(CV́.CVV) violates another property of a well-formed foot: a heavy syllable occurs in a non-head position (Weight-to-Stress Principle, Prince 1991; Hayes 1995). In this hierarchy, the trochee containing two light syllables, (CV́.CV), is the most optimal foot. The hierarchy in (3b) shows feet with the first syllable of the shape [CVV]. Like the hierarchy in (3a), a foot containing a sequence of the two syllables most different in weight, (CV́V.CV), is the least favored. However, the trochee (CV́V.CV) in (b) is better than (CV́.CVV) in (a), due to the fact that in (CV́V.CV) the heavy syllable carries the stress, while in (CV́.CVV), the heavy syllable occupies the non-head position. Also, in the hierarchy in (3b), the unmarked case is the even one, that is, the one whose syllables are identical in weight, (CV́V.CVV). Finally, the hierarchy in (3c) demonstrates trochees with the first syllable of the form [CVC]. The order of (CV́C.CVV) and (CV́C.CV), is not entirely clear, since in both trochees the second syllable deviates equally much from the first one. It is, however, clear that the best trochee also in this hierarchy is the one with syllables of equal weight, (CV́C.CVC).

I have not combined the three hierarchies into one, for two reasons. First, it is not clear whether one crucial ranking can be established between all types of trochees. Second, one conflated hierarchy is irrelevant in this study, since only certain parts of the ranking are essential. I propose, however, one type of general hierarchy: even trochees are regularly more optimal than the uneven ones (e.g. *Uneven Trochees >> *Even Trochees).

The hierarchies in (3) and the discussion above answer the question of why it is the trochee (CV́.CVV) which provides the environment for gemination: (CV́.CVV) is the *worst* possible syllabic trochee,[10] whereas all other types listed in the hierarchies in (3) are able to satisfy the foot form requirements better (i.e. the evenness in syllable weight and the resistance to heavy syllables in non-head positions). Gemination is thus an attempt to reconcile the foot-internal unevenness in the worst possible syllabic trochee – an attempt to form a better foot.

Finally, the trochaic hierarchies in (3) imply a three-way prominence distinction between the syllables [CV], [CVC], and [CVV], contrary to the traditional assumption, which categorizes [CVC] and [CVV] together as bimoraic (heavy) syllables (Karlsson 1983).[11] However, other phenomena exist in Finnish, which make a difference between the syllable structures [CVC] and [CVV], such as that of the minimal word (Harrikari 1999). Finnish exhibits a minimal word pattern in which the bimoraic [CVC], such as *[min] 'of which', fails to satisfy requirement imposed by prosodic minimality. Consequently, such word types undergo augmentation (e.g. [min-kä] 'of which'), while the bimoraic [CVV] (e.g. [maa] 'country') successfully satisfies the minimal word requirement. Thus, a prominence difference does lie between the syllable structures [CVC] and [CVV] in other areas of Finnish phonology as well. Cross-linguistically, the prominence distinction between the three syllable structures is based on sonority (Prince 1983; Laver 1994; Blevins 1995), a fact which has been supported in phonetic studies as well (e.g. Gordon 1997a, b[12]). A syllable which contains a long vowel (i.e. [CVV]) is more sonorous overall than a syllable in which the rhyme consists of a short vowel and a coda consonant (i.e. [CVC]). Assuming a positive correlation between sonority and prominence, the more sonorous a syllable is, the more prominent it is. (For a different view, based on the Dutch stress system, see Lahiri & Koreman 1988.) I incorporate the difference between [CVC] and [CVV] in my analysis of gemination. However, I do not attempt to directly formalize this distinction (e.g. through moraic structure or other structural properties, such as branching tree structure), but my analysis does reflect the difference via the trochee-oriented constraints given in (3). The upcoming analysis will show that the representations in (3) are specific enough in order to capture the patterns of gemination.

Now that I have established the concept of a more optimal trochaicity through the three hierarchies which reflect the internal balance of different types of syllabic trochees, and motivated the appearance of gemination in the trochee of the form (CV́.CVV), I will turn to the actual analysis of gemination. The formal analysis will show how the patterns governing gemination are captured by the interaction of the trochaic markedness constraints and positional faithfulness. Furthermore, I will demonstrate how the analysis implies the priority of positional faithfulness over syllable markedness.

3.2 Trochaic constraints, positional faithfulness, and syllable markedness

Gemination is a tool for achieving a more optimal syllabic trochee. Although changing the length of the syllable-initial consonant has a positive effect on the foot-internal balance, it has other consequences as well. In OT terms, gemination violates an input-output faithfulness, since the length of the word-medial consonant of the output does not correspond to that of the input. One might suggest that the inactive constraint in this pattern is the faithfulness constraint Wt-Ident (McCarthy 1995). This constraint has been employed in the literature in order to account for different types of lengthening and shortening patterns of segments. However, Wt-Ident, which requires the mora counts of output segments to be identical to that of their input correspondents, fails to account for dialectal gemination. More specifically, Wt-Ident fails to take into account the fact that gemination (i.e. (CV́C.CVV) from (CV́.CVV)) is employed in order to achieve a better prosodic balance foot-internally, as opposed to vowel lengthening (i.e. (CV́V.CVV) from (CV́.CVV)). This failure is due to the fact that both gemination and vowel lengthening violate Wt-Ident equally, since in both cases the weight of one segment is changed. The quality of the segment (i.e. consonants vs. vowels) is irrelevant in the evaluation of Wt-Ident. Furthermore, because of the dominant position of *(CV́C.CVV) over *(CV́V.CVV) (resulting from the general ranking *Uneven Trochees over *Even Trochees), the (even) trochee *(CV́V.CVV) is, in fact, more optimal than *(CV́C.CVV). Thus, the crucial question to be answered is why the attempt to balance the foot does not proceed to the more optimal trochee *(CV́V.CVV), but instead selects the less optimal one, *(CV́C.CVV). In other words, why is gemination preferred over vowel lengthening, and how can this pattern be formalized?

I suggest that the reason for employing gemination instead of vowel lengthening lies in positional faithfulness (Beckman 1998): gemination is preferred over vowel lengthening, since a change in the moraic structure of the nucleus of the stressed syllable is avoided. This is due to the privileged status of the nucleus of the stressed syllable with respect to the nuclei of unstressed syllables, a difference which can be expressed in terms of positional faithfulness. I formalize this suggestion with the faithfulness constraint given in (4).

(4) Ident-Nuc-σ_{Strs}(weight) (based on Beckman 1998)
Let β be the nucleus of a stressed syllable in the output, and α the input correspondent of β. If β is [γ weight], then α must be [γ weight].

The constraint Ident-Nuc-σ_{Strs}(weight) requires the mora count of the nucleus of a stressed syllable to be identical to its input correspondent. Consequently, any vowel lengthening or shortening is prohibited. When Ident-Nuc-σ_{Strs}(weight) is ranked below the trochaic markedness constraint *(CV́.CVV), gemination takes place, as illustrated with the partitive form /kala+μ/ 'fish' in the tableau in (5). In all subsequent tableaux, I give two representations for every output. I give a structural representation, indicating footing by parentheses, and syllabification in brackets. Long vowels are marked with two subscripted moras following the segment, and geminate consonants with one mora. The second representation is a less abstract one. In the phonetic representation, in brackets here, long segments are indicated with two graphemes in order to show the crucial syllable boundary in the middle of the geminate consonant.

(5) *(CV́.CVV) >> IDENT-NUC-σ_{Strs}(weight)

/kala+μ/	*(CV́.CVV)	IDENT-NUC-σ_{Strs}(weight)
a. (kála$_{\mu\mu}$) / [ka.laa]	*!	
☞ b. (kál$_{\mu}$a$_{\mu\mu}$) / [kal.laa]		
c. (ká$_{\mu\mu}$la$_{\mu\mu}$) / [kaa.laa]		*!

The tableau in (5) illustrates how candidate (a) loses the competition under the highest-ranked trochaic constraint *(CV́.CVV), while the output in (c) crucially fails to satisfy the faithfulness constraint IDENT-NUC-σ_{Strs}(weight). Consequently, candidate (b), the one with gemination, is correctly selected as optimal. Thus, the simple ranking of *(CV́.CVV) over IDENT-NUC-σ_{Strs}(weight) captures both the fact that fortition takes place in the foot of the shape *(CV́.CVV), and that fortition is realized as gemination.

Gemination provides evidence for the preference for geminates over long vowels, a fact which has implications with respect to the cross-linguistic typology of syllable markedness. Several cross-linguistic studies have shown evidence for the opposite ranking, in other words, for the preference for long vowels instead of gemination (McCarthy 1979; Cairns & Feinstein 1982; Hyman 1985; Hayes 1989), implying the markedness hierarchy given in (6).

(6) Syllable markedness hierarchy
CV ≻ CVV ≻ CVC = *[CVC] >> *[CVV] > *[CV]

The ranking in (6) demonstrates the idea that the existence of bimoraic [CVC] sequences entails the existence of phonemic vowel length. The intrinsic ranking in (6) suggests a preference for vowel lengthening over gemination, and that gemination would create typologically marked syllables; the typologically less marked mechanism would be vowel lengthening. Clearly, fortition realized as gemination is at odds with syllable markedness.

Optimality Theory, however, provides the appropriate tools for handling the conflict. By ranking positional faithfulness, IDENT-NUC-σ_{Strs}(weight), above the markedness scale, the desired results arise. I demonstrate the effects of the markedness scale by the two constraints, given in (7). The fact that these constraints are concerned with the properties of the second, weaker mora of the syllable is indicated by the subscripted (w).[13] (For a more detailed discussion on expressing the difference between the first, strong mora of the syllable, and the second, weak mora, see Liberman 1975; and Liberman & Prince 1977.)

(7) a. $*C_{\mu(w)}$
The second (weaker) mora of a syllable must not be associated with a consonant.
b. $*V_{\mu(w)}$[14]
The second (weaker) mora of a syllable must not be associated with a vowel.

The constraints in (7) are evaluated as follows: $*C_{\mu(w)}$ assigns a violation to every syllable containing a coda consonant, such as [CVC] and [CVVC], as codas regularly bear moras in Finnish, while $*V_{\mu(w)}$ is violated whenever a syllable includes a complex nucleus (long vowel or diphthong). Furthermore, given the cross-linguistic markedness hierarchy in (6), the two constraints in (7) must also be intrinsically ranked, $*C_{\mu(w)}$ over $*V_{\mu(w)}$. I will illustrate the crucial ranking of these markedness constraints with respect to earlier rankings, as seen in the tableau in (8). Recall that the positional faithfulness constraint Ident-Nuc-σ_{Strs}(weight) is crucially dominated by the constraint in trochaicity, *(CV́.CVV). More importantly, however, this positional faithfulness constraint must dominate syllable markedness, particularly $*C_{\mu(w)}$.

(8) *(CV́.CVV) >> Ident-Nuc-σ_{Strs}(weight) >> $*C_{\mu(w)}$ >> $*V_{\mu(w)}$[15]

/kala+μ/	*(CV́.CVV)	Ident-Nuc-σ_{Strs}(weight)	$*C_{\mu(w)}$	$*V_{\mu(w)}$
a. (kála$_{\mu\mu}$) / [ka.laa]	*!			*
☞ b. (kál$_{\mu}$a$_{\mu\mu}$) / [kal.laa]			*	*
c. (ká$_{\mu\mu}$la$_{\mu\mu}$) / [kaa.laa]		*!		**

The tableau in (8) illustrates how candidate (a) loses the competition under the highest-ranked trochaic constraint *(CV́.CVV), while the output in (c) crucially fails to satisfy the faithfulness constraint Ident-Nuc-σ_{Strs}(weight). Consequently, candidate (b), the one with gemination, is correctly selected as optimal. The crucial violations are thus identical to those of the tableau in (5). However, the evaluation in (8) shows how syllable markedness correctly fails to show its effects, due to the fact that it is dominated by positional faithfulness. The ranking *(CV́.CVV) >> Ident-Nuc-σ_{Strs}(weight) >> $*C_{\mu(w)}$ >> $*V_{\mu(w)}$ thus captures the passive role of syllable markedness, the fact that gemination is preferred over vowel lengthening in the attempt to balance the syllabic trochee (CV́.CVV), and finally, that the balancing of the foot does not proceed to the most optimal option, (CV́V.CVV).

Next, I will demonstrate how the analysis correctly predicts that gemination takes place, not only in the head foot of a prosodic word, but also in feet containing secondary stress. The evaluation of a four-syllable word,

/kumartele+μ/ 's/he nods, frequentative', in the tableau in (9), demonstrates a situation in which gemination occurs in a non-head foot.

(9)

/kumartele+μ/	*(CV́.CVV)	IDENT-NUC-σ_{Strs}(weight)	*$C_{\mu(w)}$	*$V_{\mu(w)}$
a. (kúmar)(tè$le_{\mu\mu}$) / [ku.mar.te.lee]	*!		*	*
☞ b. (kúmar)(tè$l_{\mu}e_{\mu\mu}$) / [ku.mar.tel.lee]			**	*
c. (kúmar)(t$è_{\mu\mu}le_{\mu\mu}$) / [ku.mar.tee.lee]		*!	*	**

The tableau in (9) illustrates how the candidate with gemination in the second (non-head) foot, candidate (b), is selected as optimal, since the other two candidates lose the competition under the higher-ranked constraints *(CV́.CVV) and IDENT-NUC-σ_{Strs}(weight). The ranking thus continues to make correct predictions even in cases of secondary stress.

Finally, I discuss situations in which gemination fails to take place for various reasons. First, the question arises of how to formally capture the fact that the syllabic trochee of the form (CV́.CVV) is the *only* foot type which triggers gemination. Formally, the solution lies in the crucial ranking of the rest of the trochaic constraints given in (3): all these other constraints are ranked lower than the mora markedness constraints (i.e. *(CV́.CVV) >> IDENT-NUC-σ_{Strs}(weight) >> *$C_{\mu(w)}$ >> *$V_{\mu(w)}$ >> ALL OTHER TYPES OF SYLLABIC TROCHEES). This implies the acceptability of the other types of syllabic trochees, a fact which holds even if the trochee is not optimal in the sense that the two syllables are not even, or that a heavy syllable occupies a non-head position. This result is exemplified by evaluating the genitive form of /kala/ 'fish' (which is /kala+n/) in the tableau in (10). The word consists of the syllabic trochee (CV́.CVC) which is not expected to trigger gemination, due to the low-ranking of the constraint which militates against this shape of trochee.

(10) *(CV́.CVV) >> IDENT-NUC-σ_{Strs}(weight) >> *$C_{\mu(w)}$ >> *$V_{\mu(w)}$ >> *(CV́.CVC)

/kala+n/	*(CV́.CVV)	IDENT-NUC-σ_{Strs}(weight)	*$C_{\mu(w)}$	*$V_{\mu(w)}$	*(CV́.CVC)
☞ a. (kálan) / [ka.lan]			*		*
b. (kál$_{\mu}$an) / [kal.lan]			**!		
c. (ká$_{\mu\mu}$lan) / [kaa.lan]		*!	*	*	

The evaluation in (10) demonstrates how the effects of the low-ranked constraint *(CV́.CVC) are masked, while the most faithful candidate, (a), is selected as optimal. Any change in segmental length has fatal consequences; either a violation of faithfulness (i.e. Ident-Nuc-σ_{Strs}(weight)), or an unnecessary violation of markedness (i.e. *$C_{\mu(w)}$). These results are desirable, since only the worst possible trochee, (CV́.CVV), triggers gemination.

Another reason why gemination might be blocked is the fact that gemination would occur across a foot boundary. This observation provides strong evidence for the assumption that gemination is truly connected to footing. The evaluation of the partitive singular of /omena-/ 'apple' in the tableau in (11) shows how gemination is blocked at the foot boundary. (For simplicity, I do not show the candidate with vowel lengthening, since it loses the competition under the positional faithfulness constraint Ident-Nuc-σ_{Strs}(weight).)

(11)

/omena+μ/	*(CV́.CVV)	Ident-Nuc-σ_{Strs}(weight)	*$C_{\mu(w)}$	*$V_{\mu(w)}$
☞ a. (óme)na$_{\mu\mu}$ / [o.me.naa]				*
b. (óme)n$_{\mu}$a$_{\mu\mu}$ / [o.men.naa]			*!	*

The tableau in (11) shows that the markedness constraint *$C_{\mu(w)}$ effectively ensures that gemination takes place only in contexts which would otherwise yield the undesirable trochee (CV́.CVV). Gemination is blocked in all other contexts (i.e. across foot boundaries) by *$C_{\mu(w)}$, since in those environments, *$C_{\mu(w)}$ is unnecessarily violated, as exemplified by candidate (b). Consequently, the candidate in (a), the one without gemination, is correctly predicted to be the optimal output.

I have thus demonstrated how gemination results from the interaction of the crucial trochaic constraint *(CV́.CVV) and positional faithfulness, Ident-Nuc-σ_{Strs}(weight). Furthermore, the fact that gemination creates typologically marked syllable structures is captured by the dominant position of positional faithfulness over syllable markedness. I have shown that gemination is clearly a phenomenon connected to the internal balance of one type of syllabic trochee, (CV́.CVV), since it is blocked in all other types of syllabic trochees (expressed by the lower ranking of all other trochaic constraints), as well as at foot boundaries. Interestingly, the analysis illustrates how the syllable markedness hierarchy, the effects of which are blocked in the environment of gemination, does show its power in the context in which gemination does not occur; syllable markedness functions as the mechanism which ensures that no unnecessary gemination takes place.

4. Conclusions

This paper has demonstrated that synchronically, dialectal gemination in Finnish is not a matter of a particular syllable type, but rather that of a higher prosodic unit; it is an attempt to achieve foot-internal balance. Crucially, I have shown how preference for different types of syllabic trochees can be ranked in an Optimality Theoretic way, and how the interaction of this ranking and positional faithfulness explain why gemination is triggered only by one type of syllabic trochee, (CV́.CVV), while other trochees remain unchanged. Furthermore, this ranking crucially explains why gemination is the only viable means of improving foot balance: the other plausible mechanism, vowel lengthening, changes the properties of the nucleus of the stressed syllable, an effect which is militated against by positional faithfulness.

Finally, I have illustrated how gemination, in fact, violates cross-linguistic syllable markedness by creating marked syllables, [CVC], even though less marked structures would be available through vowel lengthening (i.e. [CVV]). The analysis shows that the violation of syllable markedness is adequately accounted for by the dominant position of positional faithfulness constraint over the entire syllable markedness hierarchy.

The analysis I have presented in this paper demonstrates and formalizes a situation in which syllabic trochees appear as quantity-sensitive. This is in direct contrast to earlier assumptions that syllable weight is irrelevant in syllabic trochee systems. My analysis thus suggests that the traditional claim that only moraic trochees and iambic systems are quantity-sensitive must be reexamined.

Notes

* Thanks to Patrik Bye, John McCarthy, Steve McCartney, Jan-Ola Östman, and Lisa Selkirk for discussion on earlier versions of this paper and related matters. I am also grateful to the participants of HILP 4 at the University of Leiden (especially Stuart Davis) for helpful suggestions. All errors remain my responsibility. This research was supported by a Kone Foundation scholarship, a Young Researcher's Grant from the University of Helsinki, a grant from the Finnish Concordia Alliance, and a grant from the Academy of Finland.

1. A syllabic trochee is a left-headed foot which consists of two syllables.

2. Kager (1992, 1993), however, criticizes this generalization, and claims that systems which have traditionally been characterized as quantity-insensitive, in fact, take syllable weight into account, at least to certain extent. Thus, according to Kager, truly quantity-insensitive systems are rare.

3. Following Finnish orthography, I indicate long segments with sequences of two identical graphemes. Finnish exhibits phonemic vowel length as well as phonemic consonant length.

4. In general, dialectal gemination has been divided into three main types. The one I discuss in this paper is called "general gemination". A complete constraint-based analysis for the other types of gemination – eastern gemination and southwestern gemination – is provided in Harrikari (to appear).

5. Briefly, primary stress falls on the first syllable and secondary stress on every other syllable thereafter. However, heavy syllables not in the initial foot have the property of attracting stress from an immediately preceding light syllable. For analysis of this partly quantity-sensitive stress assignment, see Carlson (1978), Hanson and Kiparsky (1996), Alber (1997), Elenbaas (1999), McCartney (1999).

6. Stems may be affected by consonant gradation depending on the structure of the ending attached. For example, the stem /jaka-/ may appear as [jaa-], when the second syllable is closed by the singular first person ending [-n] ([jaan] 'I share' < *[ja.kan]).

7. The grapheme <ä> represents the low, front, unrounded vowel [æ].

8. The second syllable may also be of the shape [CVVC], as the data in (1) show. The two syllable structures, [CVV] and [CVVC], however, pattern identically with respect to gemination, the crucial property being the complex nucleus.

9. Kenstowicz (1994b, citing Skousen 1975), however, proposes that gemination is a phenomenon connected to the main-stressed syllable. Briefly, he suggests that the light syllable [CV] is strengthened, due the bimoraicity requirement for the first, main-stressed syllable. Simultaneously, he unnecessarily complicates the parsing pattern in Finnish, assuming the substitution of a moraic parse for the regular syllabic trochee.

10. The trochee of the shape (ĹH) is also called "resolved trochee" (Allen 1973; Dresher & Lahiri 1991; Hanson & Kiparsky 1996).

11. The bimoraicity of [CVC] in Finnish has been independently evidenced in studies in stress assignment (Carlson 1978; Hanson & Kiparsky 1996; Alber 1997; Elenbaas 1999; McCartney 1999). The structure [CVC] is able to attract secondary stress away from a [CV] syllable, even if the [CVC] syllable appears in a word-final position, as in [ús.ko.ma.tòn] 'unbelievable' from /uskomaton/ (cf. [ús.ko.mà.ton]).

12. Gordon (1997a, b), however, follows the traditional view and assumes that, in Finnish, both [CVC] and [CVV] count as equally heavy.

13. The fact that some languages assign the second, weaker mora only to particular types of consonants has been evidenced and further discussed in several studies (Hyman 1985; Zec 1988, 1995; Hayes 1989; Steriade 1990; Sherer 1994; Blevins 1995). For example, difference is made between sonorant consonants and obstruents. Thus, the mechanism which assigns a mora to coda consonants, Weight-by-Position (Hayes 1989, 1995), takes into ac-

count the quality of the segments to which moras are linked. The constraints in (7) are only more generalized forms of the same idea: instead of distinguishing between more and less sonorous consonants (i.e. sonorants vs. obstruents), the distinction is made between two more comprehensive categories of sonority, namely consonants and vowels.

14. The effects of this constraint in this analysis are similar to constraints against long vowels proposed in other studies, such as Sherer (1994, *Long-V), Rosenthall (1994, NoLongVowel), and Borowsky and Harvey (1997, *VV). My constraint is, however, more general in the sense that it prohibits diphthongs as well.

15. One reviewer suggested that the preference for gemination over vowel lengthening could be captured by ranking the faithfulness constraint Parse-$V_{\mu\mu}$ over the trochaic constraint *(CV́.CVV) as well as over the markedness constraint against long vowels, *$V_{\mu\mu}$ (the effects of which are equal to my constraint *$V_{\mu(w)}$). Consequently, long vowels of the input are maintained, while the creation of new long vowels is prohibited. This account, however, leads to typologically problematic predictions about syllable markedness. The analysis requires the constraint against geminates to be ranked *below* the constraint against long vowels. It follows that cross-linguistically, geminates should be preferred over long vowels. As showed by the syllable markedness hierarchy in (6), this is not the case.

References

Alber, B. (1997). Quantity Sensitivity as Constraint Interaction. In G. E. Booij & J. van de Weijer (Eds.), *Phonology in Progress – Progress in Phonology, HIL phonology papers 3* (pp. 1–45). The Hague: Holland Academic Graphics.

Allen, W. S. (1973). *Accent and Rhythm: Prosodic Features of Latin and Greek: A Study in Theory and Reconstruction*. Cambridge: Cambridge University Press.

Beckman, J. (1998). Positional Faithfulness. Ph.D. dissertation. Amherst: University of Massachusetts.

Blevins, J. (1995). The Syllable in Phonological Theory. In J. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 206–244). Oxford: Basil Blackwell.

Borowsky, T., & Harvey, M. (1997). Vowel-Length in Warray and Weight Identity. *Phonology, 14*, 161–175.

Cairns, C. E., & Feinstein, M. H. (1982). Markedness and the Theory of Syllable Structure. *Linguistic Inquiry, 13*, 193–225.

Carlson, L. (1978). Word Stress in Finnish. Ms., MIT.

Dresher, B. E., & Lahiri, A. (1991). The Germanic Foot: Metrical Coherence in Old English. *Linguistic Inquiry, 22*, 251–286.

Elenbaas, N. (1999). A Unified Account of Binary and Ternary Stress. Considerations from Sentani and Finnish. Ph.D. dissertation, Utrecht Institute of Linguistics, OTS.

Gordon, M. (1997a). A Phonetically-driven Account of Syllable Weight. Ms., Los Angeles, University of California.

Gordon, M. (1997b). Phonetics and Phonological Categories: the Case of Syllable Weight. Ms., Los Angeles, University of California.

Hanson, K., & Kiparsky P. (1996). A Parametric Theory of Poetic Meter. *Language, 72*, 287–335.

Harrikari, H. (1999). The Minimal Word and Augmentation in Finnish. Ms., University of Helsinki.

Harrikari, H. (to appear). From Phonology to Morphology: A Synchronic Account of Dialectal Gemination in Finnish. *Proceedings of the 11th Student Conference in Linguistics*, University of Texas, Austin. MIT Working Papers in Linguistics.

Hayes, B. (1989). Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry, 20*, 253–306.

Hayes, B. (1995). *Metrical Stress Theory. Principles and Case Studies*. The University of Chicago Press.

Holman, E. (1976). A non-segmental phonological process in Finnish: the functional unity of coda lengthening and gemination. In Fred Karlsson (Ed.), *Papers from the 3th Scandinavian Conference of Linguistics* (pp. 189–201). Turku: Academy of Finland.

Hyman, L. M. (1985). *A Theory of Phonological Weight*. Dordrecht: Foris Publications.

Jalava, A. (1971). Lounaismurteiden geminaatio. [Gemination in Southwestern Dialects]. MA Thesis, University of Turku.

Kager, R. (1992). Are there any truly quantity insensitive systems? In L. A. Buszard-Welcher, L. Lee, & W. Weigel (Eds.), *Proceedings of the BLS, 18*, 123–132.

Kager, R. (1993). Alternatives to the Iambic-Trochaic Law. *Natural Language and Linguistic Theory, 11*, 381–432.

Kangasmaa-Minn, E. (1978). *Suomen äännehistorian suuntaviivat*. [Key points of the history of Finnish phonology]. Publications of Department of Finnish language and General Linguistics, University of Turku, 7.

Karlsson, F. (1983). *Suomen kielen äänne- ja muotorakenne*. [The structure of Finnish phonology and morphology] WSOY. Juva.

Kenstowicz, M. (1994a). *Phonology in Generative Grammar*. Cambridge, MA: Blackwell Publishers.

Kenstowicz, M. (1994b). On Metrical Constituents: Unbalanced Trochees and Degenerate Feet. In J. Cole & C. Kisseberth (Eds.), *Perspectives in Phonology* (p. 51). CSLI Publications.

Lahiri, A., & Koreman, J. (1988). Syllable Weight and Quantity in Dutch. *Proceedings of WFFCL, 7*, 217–228.

Laver, J. (1994). *Principles in Phonetics*. Cambridge: Cambridge University Press.

Liberman, M. (1975). The Intonational System of English. Ph.D. dissertation, MIT.

Liberman, M., & Prince, A. (1977). On Stress and Linguistic Rhythm. *Linguistic Inquiry, 8*, 249–336.

McCarthy, J. (1979). Formal Problems in Semitic Phonology and Morphology. Ph.D. dissertation, MIT. Distributed by Indiana University Linguistics Club. Published by Garland Press, New York, 1985.

McCarthy, J. (1995). Faithfulness in Prosodic Morphology & Phonology: Rotuman Revisited. Ms., University of Massachusetts.

McCarthy, J., & Prince, A. (1986). Prosodic Morphology. Ms., University of Massachusetts, Amherst & Brandeis University.

McCarthy, J., & Prince, A. (1993). Prosodic Morphology I: Constrain Interaction and Satisfaction. Ms., University of Massachusetts, Amherst and Rutgers University.

McCartney, S. (1999). *Ternarity in Finnish Stress*. Paper presented at the 11th Student Conference in Linguistics. Austin: University of Texas.

Nahkola, K. (1987). Yleisgeminaatio. Äänteenmuutoksen synty ja vaiheet kielisysteemissä erityisesti Tampereen seudun hämäläismurteiden kannalta. [Primary Gemination. The Actuation and Implementation of a Sound Change with Special Reference to the Häme Dialects of Finnish.] *Finnish Literature Society, 457*.

Palander, M. (1987). Suomen itämurteiden erikoisgeminaatio. [Secondary gemination in the eastern Finnish dialects.] *Finnish Literature Society, 455*.

Paunonen, H. (1973). On the Primary Gemination of Finnish Dialects. *Finnisch-Ugrische Forschungen, 40*, 146–164.

Paunonen, H. (1974). Monikon genetiivin muodostus suomen kielessä I. [Formation of plural genitive in Finnish]. *Finnish Literature Society, 317*.

Prince, A. (1983). Relating to the Grid. *Linguistic Inquiry, 14*, 19–100.

Prince, A. (1991). Quantitative Consequences of Rhythmic Organization. Papers from the 26th Regional Meeting. *Chicago Linguistic Society*, Vol. 2, 355–398.

Prince, A., & Smolensky, P. (1993). Optimality Theory: Constraint Interaction in Generative Grammar. Ms., Rutgers University, New Brunswick, and University of Colorado, Boulder, RuCCS-TR-2. To appear, Cambridge, MA: MIT Press.

Rapola, M. (1961). Johdatus suomen murteisiin. [Introduction to Finnish Dialects.] *Tietolipas, 4* (2nd edition).

Rapola, M. (1966). Suomen kielen äännehistorian luennot. [Lectures on the History of Finnish Phonology.] *Finnish Literature Society, 283*.

Rosenthall, S. (1994). Vowel/Glide Alternation in a Theory of Constraint Interaction. Ph.D. dissertation, Amherst, University of Massachusetts

Sherer, T. (1994). Prosodic Phonotactics. Ph.D. dissertation, Amherst, University of Massachusetts.

Skousen, R. (1975). *Substantive Evidence in Phonology*. The Hague: Mouton.

Steriade, D. (1990). Moras and Other Slots. In *Proceedings of the First Meeting of the Formal Linguistics Society of Midamerica*, 254–280.

Zec, D. (1988). Sonority Constraints on Prosodic Structure. Ph.D. dissertation, Stanford University.

Zec, D. (1995). Sonority Constraints on Syllable Structure. *Phonology, 12*, 85–129.

Ternarity is Prosodic Word binarity*

Jay I. Rifkin
Massachusetts Institute of Technology

1. Introduction

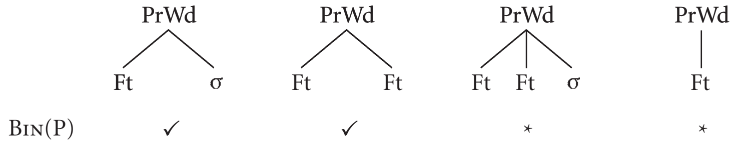
Languages such as Cayuvava and Chugach, in which stress falls on every third syllable of the word, used to have a place within a theory of foot inventories, as cases of ternary branching feet. Recent developments in metrical theory, however, have considerably reduced the space of possible feet, leaving no room for ternary iterating systems or for ternary windows for stress placement.

Extensions proposed so far to recapture these phenomena fall into two categories: constraints on foot placement and enrichments of the prosodic hierarchy. These various extensions capture the stress facts of ternary systems, but no independent evidence has been adduced for their existence, calling their status as explanations into question. In this paper, I instead pursue a more restrictive and thus stronger hypothesis, offering an account of these languages which operates within the current prosodic hierarchy and within the constraint families which interact to derive the more familiar binary language systems. Namely, I propose the following, building on the insights of Itô and Mester (1992) and Hewitt (1992): the ternary structures in ternary iterative and windowing systems are *Prosodic Words*. I further propose that Prosodic Words are constrained to this ternary shape by the violable constraint Bin(P):

(1) Bin(P): PrWds are binary, meaning that a PrWd node in a tree has exactly two daughters. One violation is incurred for each extra or missing branch.

Its effects are illustrated in (2), for a sampling of PrWd shapes:

(2)



Thus a ternary (3-syllable) unit is a prosodic word consisting of a foot plus an unparsed (meaning unfooted) syllable. By placing a further restriction on prosodic structure, namely that binarity constrains not only feet but prosodic words, we unyoke the PrWd from the lexical word. A lexical word may contain more than one PrWd and it need not be exhaustively parsed by PrWds, depending on the interaction of Bɪɴ(P) with other constraints in the grammar.

If ternary units are indeed Prosodic Words, we might expect to see two phenomena in ternary languages which we would not anticipate under previous hypotheses. First, in languages where prosodic alignment outweighs morphological alignment, we would expect to see the effects of phrasal prosody on the stress of lexical items. Second, we might expect to see Minimal Word effects somewhat akin to those noted by Itô and Mester (1992) for Japanese.

This is indeed what we find. In the course of the paper, I will present three such pieces of evidence which provide crucial evidence for the binary Prosodic Word: (a) phrasal effects on stress in Cayuvava, (b) minimal word effects on segmental processes in Cayuvava, and (c) minimal word effects on phrasal stress in Pirahã.

The paper proceeds as follows: first, I review the generation of the familiar binary rhythmic typology and show how the addition of Bɪɴ(P) to the constraint set yields ternary structures, which I use in an account of basic Cayuvava word stress. Then I turn to phrasal stress and segmental processes in Cayuvava, demonstrating that the constraint ranking developed for the word level makes the correct predictions, whereas other accounts of ternary stress do not. I then consider the other half of the ternary typology, ternary windows, accounting for word level stress in Pirahã. Again, I turn to phrasal effects in Pirahã for crucial evidence that the binary Prosodic Word account is correct, whereas previous accounts are not. In particular, I argue that the Lapse-type constraints of the sort proposed by Green and Kenstowicz (1995) and Kager (1995) can be eliminated from the universal constraint set. Lastly, I explore the further typological consequences of adding Bɪɴ(P) to the constraint set, including the prediction of quaternary windowing systems such as Palestinian Arabic.

2. Generating the binary rhythmic pattern (cf. Kager 1995)

The following three constraints determine the familiar binary typology:

(3) Parse(σ): Syllables are parsed by feet. (P&S)

(4) FtBin: Feet are binary at some level of analysis, moraic or syllabic. (P&S)

(5) All-Ft-L (AFL): Align(Ft, L, PrWd, L). The left edge of every foot coincides with the left edge of a Prosodic Word. (Left is used for concreteness. All-Ft-R also exists.) (M&P)

A sample ranking gives the following output:

(6)

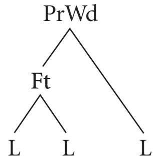
Candidates	FtBin	Parse(σ)	AFL
a. #(σσ)σσσσ...#		****...	
b. #(σσ)(σσ)(σσ)...#			******...
c. #(σσσσσσ...)#	****...		
d. #(σσσ)σσσ...#	*	***...	
e. #(σσ)σ(σσ)σ...#		**...	***...

If All-Ft-L is lowest-ranked, the iterative binary pattern emerges, as in (6b) (e.g., Maranungku, Weri, etc.). Permuting the rankings, if FtBin is lowest, unbounded feet are preferred (6c) (e.g., French, Latvian), and if Parse(σ) is lowest, a single edgemost foot is optimal (6a) (e.g., Capanahua). This three-way typology (ignoring for the moment differences in treatment of monosyllabic feet) does not directly yield ternary rhythms, but it does account for the basic stress pattern of the majority of languages, and it does allow for iterativity and edgemost windows, the crucial properties we need to capture. Thus extending this system to ternarity seems a natural way to proceed.

3. Counting to three by counting to two: Ternary rhythm with binary feet

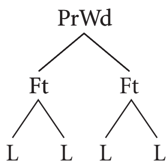
As noted above, the key observation about a ternary interval is that it can be viewed as containing a binary unit plus an unparsed syllable. If the binary unit is the foot, and if PrWd can directly dominate a syllable, the tree for a foot plus unparsed syllable is the following:

(7)

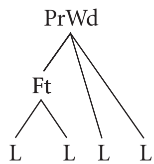


This is a perfectly reasonable option for three-syllable words given the set of constraints above (where FtBin >> Parse(σ)), but for longer words, the ternary structure is lost. Specifically, to achieve ternarity, we need to rule out the following formations:

(8)



(9)



(8) can be eliminated by ranking All-Ft-L >> Parse(σ), which demands that syllables be left unparsed by feet rather than allowing in more than one foot. With only (3)–(5), however, this results in (9) as the optimal candidate. To rule this out, we need to constrain the size of the prosodic word. We can do this in the same way we limit foot size, by introducing a constraint in the Binarity family, Bin(P), given in (1) above.

Ranking Bin(P) >> Parse(σ) gives the desired result, forcing the added syllable outside the PrWd:

(10) Ternarity is Prosodic Word binarity

			All-Ft-L	Bin(P)	Parse(σ)
☞	a.	[(σσ)σ]σ			**
	b.	[(σσ)(σσ)]	*!*		
	c.	[(σσ)σσ]		*!	**
	d.	[(σσ)]σσ		*!	**

Note: Here and below, ( ) are foot boundaries; [ ] are Prosodic Word boundaries

This establishes the basic ternary shape, but there must also be a way of allowing or blocking its propagation across the lexical item, to account for the difference between iterative and windowing systems. Allowing them is straightforward; it follows from the demands that all syllables be parsed into feet (Parse(σ)), and that the edge of a foot coincide with the edge of a prosodic word, as in (11) below:

(11)

		All-Ft-L	Bin(P)	Parse(σ)
	a. [(σσ)(σσ)]σσ	*!*		**
	b. [(σσ)][(σσ)][(σσ)]		*!**	
	c. [(σσ)(σσ)(σσ)]	*!*****	*	
	d. [(σσ)σ](σσ)σ	*!**		
☞	e. [(σσ)σ][(σσ)σ]			**

To constrain the iteration, we need a constraint on prosodic word placement of the type that restricts binary alternation, namely a member of the Align family:

(12) All-P-L: Align(PrWd, L, MWd, L). The left edge of every PrWd is aligned with the left edge of a Morphological Word.[1] As with other Align constraints, violations are assessed to intervening syllables. (L for concreteness; All-P-R also exists.)

If All-P-L >> Parse(σ), one PrWd is generated at the lexical word edge, (13b, c). If Parse(σ) >> All-P-L, PrWds iterate along the lexical string, (13a).

(13)

	All-Ft-L	All-P-L	Parse(σ)
a. [(σσ)σ][(σσ)σ]		***	**
b. [(σσ)σ]σσσ			****
c. [(σσ)(σσ)]σσ	**		**

Note that so far, we have not allowed for the possibility of PrWd recursion, that is, allowing one PrWd to contain another PrWd. This option would force us to consider candidates such as (14a–c) below:

(14)

	All-Ft-L	All-P-L	Parse(σ)
a. [(σσ)[(σσ)σ]σ		**	**
b. [[[[(σσ)σ]σ]σ]σ]			****
c. [[(σσ)(σσ)](σσ)]	******		
d. [(σσ)(σσ)](σσ)	??		

The first can be ruled out, but the second and third each fare perfectly on two constraints; thus we might expect them to be good under some ranking, if recursion is possible. Additionally, given that we are already permitting violations of Strict Layering by allowing PrWds to directly dominate syllables, we might also need to consider (14d), in cases where exhaustive parsing dominates foot alignment (Bin(P) >> Parse(σ) >> All-Ft-L). Such structures should always win if Parse(σ) is highly ranked, as we see below:[2]

(15)

	Candidates	All-P-L	Parse(σ)	Bin(P)	All-Ft-L
	a. [(σσ)σ][(σσ)σ]	*!**	**		
	b. [(σσ)(σσ)]σσ		*!*		**
	c. [(σσ)(σσ)(σσ)]			*!	******
(14)	d. ☞ [(σσ)(σσ)](σσ)				??
	e. [(σσ)(σσ)][(σσ)]	*!***		*	**

In this paper I will argue that the stress systems we observe can be accounted for without allowing PrWd recursion (14b, c), but extending this account to quaternary systems will require us to consider non-exhaustive foot parsing (14d). Let us then pursue the strongest hypothesis possible, by making the following restrictive assumptions about metrical structure:[3]

(16) Stress is assigned to foot heads, or more precisely, foot heads project a mark on the next grid layer, and prominence on the grid is interpreted as stress.

This precludes the possibility of headless feet or prosodic words. It also precludes the possibility to assigning stress to a syllable which is not parsed by any foot.

(17) Prosodic constituents do not recurse.

Here, the important part of this claim will be that Feet and Prosodic Words do not recurse. I have no evidence bearing on the recursion properties of higher level constituents such as Phonological Phrases.

These assumptions result in a reasonable set of candidates for ternary structure, but since these are claims about the universal candidate set, adopting (17) may require a rethinking of traditional views on cyclic-looking affixation, in which it has been supposed that PrWds do recurse to form larger units, e.g.,

(18) [[[help]less]ness].

However, the typological predictions made here do seem consistent with the structures assigned to other sorts of affixation and to compounds.

Given the provisions above, we see that the constraints in (1), (3)–(5), and (12) establish the possibility of ternary constituents consisting of a foot plus a light syllable. The rankings in (13) determine the number of PrWds generated, (1) restricts their size, and (3)–(5) determine the structure within each PrWd. We turn now to the two stress patterns discussed so far, ternary iteration and ternary windows, to demonstrate the viability of this system and further motivate the ternary structure proposed above.

4. A ternary iterative stress system – Cayuvava

Cayuvava is a quantity-insensitive language which was spoken in the Northern Bolivian lowlands. It is almost certainly now extinct.

4.1 Regular stress *(data from Key 1961, 1967; Levin 1988)*

(19)
a. dá.pa 'canoe'
b. tó.mo.ho 'small water container'
c. a.rí.po.ro 'he already turned around'
d. a.ri.pí.ri.to 'already planted'
e. á.ri.hi.hí.be.e 'I have already put the top on'
f. ma.rá.ha.ha.é.i.ki 'their blankets'
g. i.ki.tá.pa.re.ré.pe.ha 'the water is clean'
h. tʃá.a.di.ró.bo.βu.rú.ru.ce 'ninety-nine (first digit)'
Note: Cayuvava has no long vowels. VV is disyllabic.

Note the "double upbeat." The first two syllables of a 5- or 8-syllable word are not stressed.

The stress pattern of Cayuvava can be summarized as follows: stress the initial syllable in disyllables; in longer words, stress every third syllable from the end. There are no monosyllabic words. There is no apparent primary/secondary stress distinction.

Invariant initial stress in the ternary groupings suggests that the following constraint is undominated:

(19) TROCH: Foot Form is trochaic.

To generate dactyls, we need ALL-FT-L >> PARSE(σ), which demands just one foot per PrWd, and BIN(P) >> PARSE(σ), which forces each PrWd to be larger than just one foot (and FTBIN undominated, to block the generation of just one very big foot). This is demonstrated on a short word below:

(20)

a.rí.po.ro[4]	FTBIN	AFL	BIN(P)	PARSE(σ)
a. [(á.ri)(pó.ro)]		*!*		
b. [(á.ri)po.ro]			*!	**
c. [a(rí.po)ro)]		*!	*	**
d. a.ri[(po.ro)]			*!	**
☞ e. a[(rí.po)ro]				**

Ternary rhythm iterates, so PARSE(σ) >> ALL-P-R. These rankings give us the correct result, as shown for a sampling of candidates for a longer word below:

(21) Tableau for /arihirataka/ 'we are already old'

á.ri.hi.rá.ta.ka	FTBIN	AFL	BIN(P)	PARSE(σ)	ALL-P-R
a. [(á.ri)(hí.ra)(tá.ka)]		*!*****	*		
b. [(á.ri.hi)(rá.ta.ka)]	*!*	***			
c. [(á.ri)hi(rá.ta)ka]		*!**	*	**	
☞ d. [(á.ri)hi][(rá.ta)ka]				**	***

In addition, we want to be sure that this ranking yields the correct result for the double upbeat:

(22) Double upbeat

	FTBIN	AFL	BIN(P)	PARSE(σ)
☞ a. #σσ[...				**
b. #[(σ)σ][(...	*!			*
c. #[(σ)(σ)][...	*!*	*		
d. #[(σσ)][...			*!	

Note that in (22a, b), we also have evidence for the crucial ranking of FTBIN >> PARSE(σ), as degenerate footing is not allowed even to achieve a more complete parse.

However, for disyllabic words, the candidate foot structures are the same as in (22), but a single initial stress wins out, meaning either (22b) or (22d) must win out. In order for that to happen, the following familiar constraint must be at work:

(23) LX~PR: A lexical word contains a PrWd. (P&S)

LX~PR dominates either FTBIN or BIN(P). To determine which, we might wish to turn to the absence of monosyllables in Cayuvava,

(24)

σ	FTBIN/BIN(P)	LX~PR	BIN(P)/FTBIN
a. Ø		*	
b. [(σ)]	*!		*

but the situation is symmetrical. One dominates LX~PR, while the other is dominated by LX~PR.[5]

In sum then, constraint rankings in Cayuvava are as follows:

(25) Parse(Segment), FtForm
/ \
All-Ft-L, FtBin,
\ /
Bin(P)
|
Parse(σ)
|
All-P-R,

plus some morpheme-specific constraints to account for irregular stress, which space constraints prevent us from pursuing here.

4.2 Phrases and utterances in Cayuvava

Thus far we have been considering words in isolation, but for Cayuvava, this is just a subcase of the phrasal stress pattern. In Cayuvava, Key (1967) reports that the ternary stress pattern can extend throughout an utterance, or more frequently, throughout a phrase he refers to as the stress contour word, "a phrase or group of words bound together by a distinctive intonation-stress pattern of a series of tri-syllabic strong stresses. It is composed of more than one tri-syllable group and more than one word" (p. 21). That is, the three-syllable iteration does not restart its count at word boundaries; rather, it continues across the entire stress contour domain. The intonational contour consists of a gradual pitch rise across the domain, culminating at the final stress, and followed by one of the following: "falling pitch, indicating utterance final; extended, indicating temporary terminal; and rising, indicating interrogative" (p. 20). Levin (1988) claims that these pitch changes diagnose the presence of the greater prominence of the final stressed syllable in a lexical word despite the lack of additional stress. Now that we have decoupled the lexical and the prosodic words, however, we can see that this change in pitch is just the very common process which applies following the main stress of the final prosodic word in a phrase or utterance (as in English, for example).

This provides some evidence that these ternary units are Prosodic Words, rather than subconstituents or nonconstituents. The usual prosodic word (the one generally assumed in the literature), which is coextensive with the lexical stem, plays no role here; its boundaries are apparently constantly violated and its stress pattern can vary depending on the stress count of the word that follows. This simply follows from the demands of prosodic alignment outweighing those of prosody-morphology alignment in Cayuvava. Thus, AL(P, PHON, R) >> ALL-P-R, where

(26) AL(P, PHON, R): ALIGN(PrWd, R, Phonological Phrase, R)

This sort of prosodic alignment is familiar from the work of Nespor and Vogel (1986), McCarthy and Prince (1993b), and others.

4.3 Interactions between segmental process and prosody as evidence for Prosodic Word binarity

4.3.1 *A condition on vowel deletion*

Returning now to the individual lexical item, we can uncover two more pieces of evidence in favor of an iterating prosodic word account for Cayuvava. Hiatus is allowed in Cayuvava quite freely, allowing even four vowels in sequence, as in [βádačaóai], 'my younger brother.' However, it is apparently dispreferred at morpheme junctures, and the weaker of two vowels will generally delete, or one copy will delete, if the vowels are identical.[6] Key gives the vowel strength hierarchy as i < e < a < u, o, weakest to strongest.[7] Thus,

(27) a. /ki/ + /učãrãhi/ → /kučãrãhi/ 'when he is already coming'
b. /apa/ + /uaβa/ → /apuaβa/ 'your body' (from Key 1967: 17)

However, there is a notable exception to this generalization; it does not occur when the stem plus prefix form a three-syllable word. Key states the generalization as follows:

(28) "The weakest vowel is apparently /i/ which is the first to be lost from any cluster.
Further, unless a stress pattern permits three vowels, a cluster of /i/ plus /i/ will reduce to a single /i/." (p. 18)

Thus, when deleting the vowel would result in a less optimal prosodic word than keeping it around, deletion is blocked. Thus,

(29) a. /ki/ + /iču/ → /kíiču/ 'right here', but
b. /bečosi/ + /ihi/ → /bečósihi/ 'their spit' (Key 1967: 18)

This follows directly if we rank *i+i below BIN(P).[8]

(30) Optimal PrWd formation

		Bin(P)	*i+i
☞	a. [(kí.i)ču]		*
	b. [(kí.ču)]	*!	
	c. be.čo[(sí.i)hi]		*!
☞	d. be[(čó.si)hi]		

While this phenomena falls out quite naturally from a theory containing Bin(P), which predicts the three-syllable word to be optimal in Cayuvava, it is mysterious on weak local parsing accounts of ternarity such as Hayes (1995) or Kager (1995), which drive ternarity via foot repulsion (*FtFt). According to such accounts, (30b) is the preferred candidate; it eliminates hiatus and has its foot perfectly aligned with both edges; foot repulsion is not at play here since there is only one foot:

(31) Weak local parsing makes the wrong prediction, regardless of rule order:[9]

a. ki + iču $\xrightarrow{\text{Deletion}}$ kiču $\xrightarrow{\text{Footing}}$ (kiču)

b. ki + iču $\xrightarrow{\text{Footing}}$ (kii)ču $\xrightarrow{\text{Deletion}}$ (ki)ču

Likewise, *Lapse accounts, such as those in Green and Kenstowicz (1995) or Kager (1995), should also prefer (30b) over (30a), since the constraint against lapses is not at play for either candidate. For example, the ranking given in Green and Kenstowicz (1995) for Cayuvava yields the following:

(32) Lapse makes the wrong prediction:

/ki/ + /iču/	Lapse	Main-Stress-L	Align-Ft-R	*i+i
a. [(kí.i)ču]			**	*
☜ b. [(kí.ču)]				

Unfortunately, Key has not provided us with rich enough texts to see clear examples of this phenomenon at work in a longer utterance. However, perusing his morphological examples yields several other word-level instances, among them:

(33) a. /ji/ + /iki/ = /jiiki/ 'with them' (Key 67)
b. /ji/ + /iču/ = /jiiču/ 'where is'
vs.
c. /ki/ + /iasi/ = /kiasi/ 'the man' (Key 64)

and plausibly the following, although we lack the morphological details to be certain:[10]

(34)	a.	/ˈkiibu/	‘the ditch, gulley [sic]’	(Key 64)
	b.	/ˈkiiheˈnarãsa/	‘the orange skin’ (cf. *naranja*, Spanish ‘orange’)	
	vs.			
	c.	/mei/ + /iasi/ = /meiasi/	‘the men’	(Key 50)

Coupled with the generalization Key states, these indicate that this process is not limited to a single word or morpheme and thus may be taken as evidence that the tri-syllabic binary prosodic word is acting as the minimal word in Cayuvava, as we would expect.[11]

4.3.2 *Prosodic conditioning of semivowel surface forms*

Further evidence of the prosodic word as constituent comes from the interaction of stress and semivowel surface forms in Cayuvava. Key (1961) states that /i/ and /u/ are semivowels, with consonant [y, w] and vowel allophones [i, u]. Their distribution depends on the prosody of the utterance. Key states the generalization as follows:

(35) When the antepenultimate strong stress occurs with the semivowels u and i, whether these function as vowels or consonants depends on the number of vowel nuclei following the strong stress. That is, if there are three other vowels before the end of the utterance, then the consonant form will occur, whereas if there are only two other vowels, then the semivowel i or u will occur within the stress group functioning as a vowel (to illustrate: ˈSVCV > ˈVVCV, whereas ˈSVVCV > ˈCVVCV). (p. 150)

(36) Key’s examples include:

a.	ˈiasi	[ˈiasi]	‘man’	
b.	ˈiaroho	[ˈyaroho]	‘stone’	
c.	kaˈiuaβa	[kaˈyuaβa]	‘Cayuvava language’	
d.	ˈuauaha	[ˈwawaha]	‘want’	(p. 144)
e.	uïˈrïuïrï	[wïˈrïwïrï]	‘iguana’	
f.	ˈuarie	[ˈwariɛ]	‘kill’	(p. 146)

From the stress pattern and Key’s above claim, we can see that in (36), /i/ is a vowel and thus a syllable nucleus, but in (36a–c) it is a consonant. In (1d), /u/ is a vowel, but in (1e–f), it is realized as a consonant, as Key transcribes. Again, we see the optimal prosodic word emerging. Even in cases such as (1c), which are larger than the minimal PrWd, realizing /i/ as a consonant improves parsing (over realization as a vowel), reducing the number of unfooted syllables. In (1e), both relevant candidates have two unfooted syllables, so the decision is made by AL(P, PHON, R). Semivowel realization thus provides additional justification both for the constraint ranking proposed above, and for the activity of the lower-ranking constraints in Cayuvava grammar:[12,13]

(37)

			BIN(P)	PARSE(σ)	AL(P,PHON,R)
☞	a1.	[(ˈia)si]		*	
	a2.	[(ˈyasi)]	*!		
	c1.	kai[(ˈua)βa]		***!	
☞	c2.	ka[(ˈyua)βa]		**	
	d1.	[(ˈuï)rï][(ˈuï)rï]		**	*!**
☞	d2.	wï[(ˈrïwï)rï]		**	

Having amassed evidence for the role of the binary prosodic word in an iterative system, we now turn to considering a ternary window for stress.

5. A three-syllable window for stress – Pirahã

5.1 Word-level stress in Pirahã

(38) Representative data (from Everett 1988)
Boldface and underlined = stressed

a.	.**ʔa**.ba.gi.	'toucan'
b.	.ʔa.ba.**pa**.	'Amapa' (city name)
c.	.bií.**sai**.	'red'
d.	.ho.aa.**gai**.	species of fruit
e.	.pia.hao.gi.so.**ai**.pi.	'cooking banana'
f.	.pii.kao.**bii**.gá.há.	'was certainly raining'
g.	.ʔia.bí.ka.bi.**ka**.bi.	'proper name'
h.	.káo.bii.gá.**bai**.	'almost falling'
i.	.ka.pii.ga.ka.ga.kai.ka.báo.**báo**.	'when finished doing paper words' (studying)

Everett (1988) presents data which support the following characterization of stress in Pirahã: Stress falls on the strongest (most prominent) of the final three syllables in a word. Strongest is evaluated by the hierarchy:

(39) KVV > GVV > VV > KV > GV, where K, G consonants, K = [+cons, –voice], G = [+cons, +voice]. (K is the notation in Hayes 1995)

If two candidates are equally strong, the rightmost wins.

Lacking a way to directly specify a three-syllable "window" at the edge of the word, Everett suggests the following approach:

(40) Everett's (1988) analysis:
  a. Mark the final syllable extrametrical, and form an unheaded foot on the antepenult and penult.
  b. Annul extrametricality and adjoin the final syllable to the foot.
  c. Project all footed syllables to a grid, weighting each syllable by its place in the strength hierarchy.
  d. Apply End Rule Right.

This approach captures two important generalizations:

(41) The head of the foot (assuming that the head of the foot will be the stressed syllable) is not determined by foot form, rather by prominence.

(42) Only one ternary unit is created.

Using the constraints described here, we can capture the same generalizations without recourse to extrametricality or its annulment. Again assuming undominated PARSE(Segment), we limit metrical structure to only one PrWd by ranking ALL-P-R >> PARSE(σ), as in (13). Only one foot is allowed per PrWd, so ALL-FT-R >> PARSE(σ) as well. From the data presented, it is not actually possible to determine to which PrWd edge feet align, but this turns out not to matter here, as we will see. I have simply chosen ALL-Ft-R to be explicit. Ranking BIN(P) >> PARSE(σ), we complete the restriction of the "window" to a ternary unit, consisting of a single disyllabic foot plus an uparsed syllable. In contrast with Cayuvava, Pirahã does allow monosyllables, so LX~PR must dominate both FTBIN and BIN(P) (see tableau (24)).

(43) Monosyllabic words in Pirahã (Everett 1986, 1987)

a.	go	'what'	e.	pi	'thorn'
b.	hói	'one'	f.	pií	'water'
c.	hoí	'two'	g.	soí	'leather'
d.	kao	'mouth'	h.	ti	1st person singular

Once the "window" has been created, stress must be assigned within. As Hayes (1995) suggests, this is accomplished by comparing prominence.[14] Prominence in Pirahã can be diagonalized into two separate weights: vowel weight and onset weight (Prince, class lectures 1995). Expressing this in constraint form,

(44) PkProm(Vowel): Assign the prominence peak to the syllable with the greatest vowel weight in PrWd.

(45) PkProm(Onset): Assign the prominence peak to the syllable with the greatest onset weight (strength) in PrWd.

To achieve the strength hierarchy in (39), PkProm(Vowel) >> PkProm(Onset). Add

(46) RtMost: Stress the rightmost syllable in PrWd.

and PkProm(Vowel)>> PkProm(Onset)>> RtMost straightforwardly awards ties on PkProm to the rightmost candidate still in the running:[15]

(47) paó.hoa.**hai**. 'anaconda' (Everett 1985)

/paóhoahai/	PkProm(V)	PkProm(O)	RtMost
a. **paó**.hoa.hai			*!*
b. paó.**hoa**.hai			*!
☞ c. paó.hoa.**hai**			

The interaction of these constraints wholly determines the placement of stress (with the exception of words with a nominalizing suffix, as we will see below). Foot headedness and foot alignment constraints do not force violation of the prominence hierarchy and must consequently be ranked below the prominence constraints. In Pirahã, then, foot structure is malleable under the demands of stress assignment. Thus we cannot determine which edge of PrWd the foot aligns to, as the foot and the internal structure of PrWd will simply deform to meet the demands of prominence.

5.2 Single stress phrases as evidence for the binary Prosodic Word

Noun+Verb and Noun+Adjective form phrasal units in Pirahã. The second word receives greater stress than the first:

(48) /ʔapípai haáobá/ → ʔa.pí.ˌ**pa**. ho.ˈ**ao**.bá.[16]

Let us now consider the stress pattern for a phrase in which the final word contains only two syllables, not three. On the model of Cayuvava, we might well expect that the mandated ternary structure of the prosodic word would cause it to spill over into the next word. And that is indeed what happens.

(49) Single stress phrases (from Everett 1988)

a. baó.sa. bií.sai. cf. baó.saí. in isolation
'cloth' 'red'

b. ʔa.ba.pa. go.gi. cf. go.gi. in isolation
city name 'where'

c. báa.gi.so. go.gi. cf. báa.gi.so. in isolation
'many' 'where'

vs.

d. ʔa.pa.pá. ii.tá.há.
'head' 'hurts'

Everett names such instances as cases of stress shift, but in fact, in both his analysis and the account above, they follow straightforwardly from the construction of suprasyllabic prosodic units. Just as in Cayuvava, word boundaries are thrust aside in favor of forming a better PrWd by satisfying Bin(P):

(50) Single stress phrase

		/báagiso gogi/	Bin(P)	FtBin
☞	a.	báa.gi.[(so. go.)gi.]		
	b.	[(**báa**.gi.)so.] [(go.**gi**.)]	*!	
	c.	[(**báa**.gi.)so.] [(go.)(**gi**.)]		*!
	d.	[báa.gi.(**so.** go.)gi.]	*!*	

(51) Double stress phrase

		/ʔapapaí ʔiitáhá/	Bin(P)	FtBin
☞	a.	[ʔa.(pa.pa.)] [(ii.tá.)ha]		
	b.	[ʔa.pa.pa. (ii.tá.)ha]	*!**	

So far this is just like Cayuvava phrasal stress. However, in contrast to Cayuvava, Pirahã does not allow the creation of another PrWd to parse some of the remaining syllables in the leftmost word (even if there are three or more syllables remaining). But this follows directly from the difference between Pirahã and Cayuvava, namely that All-P-R >> Parse(σ)! Since every PrWd must align to the right edge of an MWd, if the right edge of the MWd is contained in another PrWd, no new PrWd can be erected.[17,18]

In addition to providing crucial evidence for this proposal, this data also allows us to refine our constraint ranking, adding Bin(P) >> All-F-L. Here we can see the effects of the left edge of the PrWd, so we know that even when stress falls on the final syllable, the PrWd is still binary. If All-Ft-L >> Bin(P), then the PrWd in (50) would consist of just a foot, and the noun would form a PrWd on its own. Therefore Bin(P) at least outranks All-F-L; we still cannot crucially rank Bin(P) and All-Ft-R.

In sharp contrast to the naturalness of the Bin(P) account, the data given above does not accord with any of the other analyses of ternary windows we have considered so far. Lapse, *FtFt, and weak local parsing accounts fail again here for the same reason they did for Cayuvava: these theories make no predictions about minimal prosodic structure above the foot level. Therefore, they have no way of determining which phrases constitute a single domain for stress assignment, and which constitute two domains (cf. (49c) vs. (49d)). For example, using the constraint ranking given for Pirahã in Green and Kenstowicz (1995) yields the following for (49c) and (49d):[19]

(52) Lapse makes wrong prediction for Pirahã phrases

/báagiso gogi/	Lapse	PkProm	RtMost	FtForm
a. [báa.(gi.so.)] [(go.**gi**.)]		*!		*
☞ b. [(**báa**.gi.)so.] [(go.**gi**.)]			**	*
c. [(báa.gi.)(so. go.)gi.]		*!	**	

(53)

/ʔapapaí ʔiitáhá/	Lapse	PkProm	RtMost	FtForm
a. [ʔa.(pa.**pa**.)] [(ii.tá.)ha]			**	*!
☞ b. [ʔa.(pa.pa.) (ii.tá.)ha]			**	

Without a Bin(P) constraint, nothing rules out the disyllabic prosodic word *go.**gi**.* in (52b). Kager's (1995) version of the Lapse model also fails to generate this data, as the individual prosodic words entirely satisfy Parse-2. Pirahã thus constitutes a second strike against the productivity of the Lapse constraint.

Green's (1997) binary cola might be able account for this data in the same way I have done, only enforcing colon binarity rather than word binarity. If we relax the identity he assumes between morphological and prosodic words, colon binarity would force the prosodic word containing the colon to expand with the colon into the next lexical word. But given the successful account above, we can see that this would be simply redundant, adding an extra layer of structure without empirical motivation.

5.3 Nominalization

(54) Nominalization in Pirahã is achieved by adding the suffix *-sai*.

a. ʔoi.boi.**bií**.sai. (Everett 51, 52, & 54)

b. **ʔi**.bi.sai.

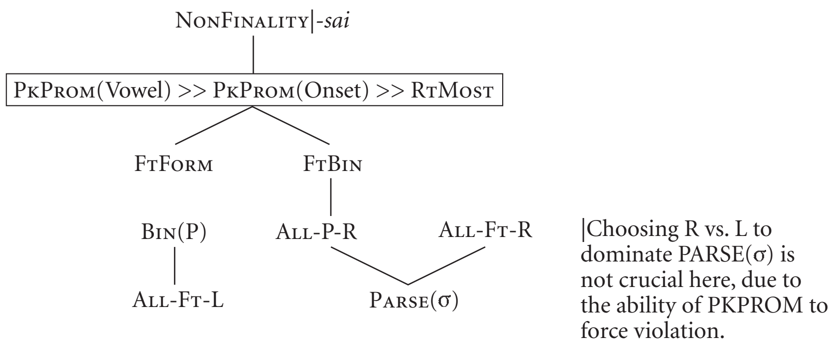
c. ko.**hoi**.bií.sai.

This suffixation poses two challenges: it shifts the placement of the "window" to the end of the word, but it cannot receive stress, no matter how it compares in prominence to the other syllables in the window. Our current model does not predict this. Because All-P-R aligns PrWds with MWds, we might expect the "window" to move upon suffixation, but we do not anticipate the peripheral extrametricality. We could imagine a constraint barring *-sai* from bearing stress, but Everett notes that it "*can* (only) be stressed when not peripheral, i.e., nonadjacent to pause" (227). This is captured by a lexical restriction on a familiar constraint, NonFinality, which we may extend to the phonological phrase:

(55) NonFinality: No Head of PhonPhr is final in PhonPhr.
(modified from P&S)

This morpheme-specific constraint then dominates the PkProm hierarchy, preventing stress assignment to phrase-final *-sai*.

(56) Final constraint ranking for Pirahã



The constraint conflicts which generate the basic ternary typology have thus been shown to be productive iterative and windowing systems. Unlike previous analyses, I have not yet been forced to rely on optional rules (weak local parsing), to increase the number of basic foot types (amphibrachs), or to propose constraints outside the currently motivated constraint families.

6. Grammar without Lapse

Positing Bin(P) accounts for the facts in ternary languages and permits a unified explanation of the stress typology. In doing so, it subsumes much of the work of Lapse. However, Lapse still performs one function in the grammar outside the scope of Bin(P). It militates against unbounded feet (greater than binary); Bin(P) does not.

However, Prince (1985) and Walker (1996) point out that in contrast to the overwhelming non-stress evidence for bounded feet, there is no independent evidence whatsoever for unbounded feet. Further, unbounded feet redundantly perform an edge-seeking function already subsumed by Align. Lastly, they argue that the stress data for which unbounded feet were originally posited can be better accounted for by other means.

If we take these findings seriously, they suggest the following proposal: Eliminate unbounded feet from the possible set of licit representations. Gen builds maximally binary feet. Applying this restriction on Gen to the candidate set in (6), we see that it does no harm at all to the typology; words in (6c, d) have the same stress pattern as words in (6a). With this restriction in place, the remaining need for Lapse disappears, and it can be eliminated from the universal constraint set.

7. Further typological consequences of this approach

We have seen that the binary PrWd provides a good account of ternary stress languages, but adding Bin(P) to the universal constraint set yields other typological predictions upon constraint reranking, and we should ask whether these predictions are borne out. Unfortunately, a complete reckoning of the typology would take us beyond the space provided; a single example and hints at others will have to suffice.

7.1 A quaternary window for stress

Ranking All-Ft-L lowest in (10) (repeated below as (57)) yields (57b) as optimal:

(57) Tableau from (10)

			Bin(P)	Parse(σ)	All-Ft-L
	a.	(σσ)σ]σ		*!*	
☞	b.	[(σσ)(σσ)]			**
	c.	[(σσ)σσ]	*!	**	
	d.	[(σσ)]σσ	*!	**	

Ranking BIN(P) and ALL-P-R over PARSE(σ) enforces a PrWd window at the word edge, two feet in size. This is a quaternary window for stress, analogous to the ternary window. If the non-edgemost foot is stressed, this window will yield effects very similar to the controversial device of foot extrametricality, proposed to account for Palestinian Arabic stress, among others.

7.1.1 *Palestinian Arabic*

(58) Representative data (from Kenstowicz 1981, 1983)

a.	ˈʃad͡ʒaratun (Classical)	‘a tree’
b.	ʕalˈlamato	‘she taught him’
c.	ˈyirʃi	‘bribe (3M.SG.IMPERF.)’
d.	bakaˈritna	‘our cow’
e.	ˈkatabna	‘we wrote’
f.	ˈbaːrako	‘he blessed him’
g.	ˈʔana	‘I’
h.	ˈkatabu	‘they wrote’

Here we abstract away from word-final effects (word-final heavies count as light; and superheavies are stressed) as they are peripheral to our main concern.

The generalization to be captured is as follows: moraic trochees iterate across the string, aligned to the left edge. If there is a stray syllable at the right, stress is assigned to the last foot; if not, stress is assigned to the penultimate foot. But this is equivalent to saying, parse *everything*, violating FTBIN, and simply stress the penultimate foot:

(59) PARSE(σ) >> FTBIN

The penultimate foot is just the first foot in a word-final binary PrWd, a quaternary window:

(60) BIN(P), PARSE(σ) >> ALL-FT-L

This plus a constraint placing main stress: ALIGN(Head(PrWd), L, PrWd, L) gives us the stress pattern of Palestinian Arabic:

(61)

	/bakaritna/	BIN(P)	PARSE(σ)	ALL-FT-L	FTBIN
☞	a. (ba.ka.)[(ˈrit.)(na)]			*,??	*
	b. [(ba.ka.)(ˈrit.)(na)]	*!		*,**	*
	c. (ba.ka.)[(ˈrit.)na]		*!	,??	
	d. ba.ka.[(ˈrit.)(na)]		**!	*	*

Note that not all feet are contained in PrWds! We had been tacitly assuming complete foot parsing, but now we see that this must really be due to the action of another parsing constraint:

(62) Parse(Ft): All Feet are parsed by Prosodic Words.

This accords with other recent work on Strict Layering violations, e.g. Hagstrom (1995) on unparsed moras, and Itô and Mester (1992) on unparsed syllables.

In addition, we must rely on a principle of phonetic interpretation to ensure that these feet are not stressed, but there are other facts about vowels in Arabic which give evidence for these foot boundaries.

7.2 Quaternary iterative stress

If we keep the ranking in (57) but instead choose Parse(σ) >> All-P-L, we get iteration of binary PrWds consisting of exactly two binary feet. This is the Hungarian reported in Hammond (1987):

(63) [(lɛ́gmɛ)(gvɛ̀stɛ)][(gɛ̂thɛ)(tɛ̀tlɛ)][(nɛ̂bbɛk)(nɛ̀k)] 'to those least bribable'
ɛ́ = primary stress; ɛ̂ = secondary stress; ɛ̀ = tertiary stress

However, I have been unable to replicate Hammond's data with native Hungarian speakers, leading me to conclude that Hungarian stress is not a clear example of a quaternary iterative stress system. Hopefully, a more solid example can be found.[20]

Note in passing that quaternary systems do not arise from Lapse-type constraint systems. Such systems need some other account of quaternary stress. But Bin(P) provides a unified explanation for binary, ternary, and quaternary systems.

7.3 Binary patterns

Ranking Bin(P) below Parse(σ) and below All-P-R forces maximal parsing of syllables into feet, but concurrently demands that there be only one well-aligned PrWd. This yields the familiar array of binary rhythmic patterns.

8. Conclusion

The refinement of the theory of prosody developed here provides a straightforward and unified account of ternary iterative stress systems and ternary windows for stress. The activity of Bin(P) in the grammar also provides an account for the phrasal stress facts and minimal word effects in these languages, facts which do not follow from previous accounts of ternary stress. Reranking Bin(P) extends the theory to cover quaternary stress systems, and coupled with a unversally obeyed restriction on foot structure, Bin(P) completely subsumes the work of Lapse.

Notes

* I would like to thank Michael Kenstowicz, Alan Prince, Cheryl Zoll, and the participants at HILP 4, the 1997 Generals Workshop at MIT, the MIT/UMass Phonology Workshop, and the MIT Phonology Circle for comments on earlier versions of this work. This work was supported in part by a National Science Foundation Graduate Research Fellowship.

1. At this point, the exact category to which PrWd aligns is not clear. I have chosen MWd for explicitness, but it might be another morphological category such as Stem, and we will see for Cayuvava, that alignment to a higher level phonological category such as Phonological Word will also play a role. The choice might also be language-particular.

2. Presumably such feet would be incorporated into the next level of metrical structure above the PrWd, allowing them to be prosodified and stressed.

3. Later in the paper I will also consider the additional restriction that feet cannot be larger than binary, in light of current thinking about ternarity and the lack of crucial evidence for unbounded feet.

4. Here and below, I list the output form in this cell rather than the input. Since we will be concerned mainly with stress, the input can be considered to be this form without the stresses, unless otherwise noted.

5. The probable resolution is to find another constraint besides Lx∼Pr which obtains here.

6. The data Key gives to demonstrate this actually illustrate a somewhat more restricted phenomenon, familiar cross-linguistically: In hiatus (at morpheme junctures), the first vowel will delete if it is weaker than the second. /i/ acts exceptionally, yielding to the stronger vowel regardless of position, again a familiar cross-linguistic phenomenon.

7. /ï/, /æ/, and /ɔ/ do not participate in this strength hierarchy, neither triggering nor undergoing deletion.

8. I do not want to claim that *i+i is the correct instantiation of the processes involved here; it is simply a stand-in which allows me to capture the relevant fact.

9. Hayes proposes a rule of final syllable extrametricality for Cayuvava, revoked when footing disyllabic words.

10. Possibly also: /mei/ + ?*tooth*? = /meisi/ 'teeth,' which would be an example of deletion improving parsing.

11. If the optimal prosodic word is functioning as the minimal word in Cayuvava, then why are disyllabic inputs not repaired? Key lists no processes of epenthesis; therefore it is reasonable to suppose that epenthesis is simply not an option in Cayuvava (Dep > Bin(P)/FtBin).

12. Key also lists one other possibly relevant form in his discussion of this phenomena (1961:50):

aNjaNˈiuhue 'carry in hand'

Note: N is not a nasal consonant, but rather Key's transcription of nasalization of the vowel and consonant that precede and follow the N.
Here, /i/ is realized as a consonant, but it seems that parsing would be more optimal if it were a syllable nucleus, since this would result in two prosodic words, [ˈaNjaNi][ˈuhue], and leave only two syllables unparsed, rather than three. We may be seeing here the result of interaction between semivowel realization and the vowel deletion process discussed above, but unfortunately, Key does not give the morphology of this word.

13. Key also lists one case of variation in the interaction of semivowels and stress: /ioene/ = [ˈyoɛnɛ ∼ iˈoɛnɛ] (1961:146).

14. Everett (1988) analyzes the prominence hierarchy as a hierarchy of weight contrasts.

15. The constraints on prominence apply within the domain of the *Prosodic Word*, which differs here from the lexical word. Thus, even though these prominence constraints are undominated, they cannot push stress beyond the bounds of the window. So far, this is just stipulative, but it makes the following strong prediction: Prominence is a PrWd-level phenomenon; one should not find prominence effects which extend over multiple PrWds. If such phenomena exist, this understanding of prominence will have to be revised.

16. "In noun+adjective and noun+verb sequences, the final vowel of the noun deletes (when preceded by another vowel), as does an initial glottal occlusive on the following adjective or verb." (Everett, p. 221) Stress placement depends only on the output of this deletion; thus stress could be on the final syllable of a noun in isolation but on the antepenult in a phrase (or vice versa):

ʔáa.pa.<u>**haí**</u>. but <u>ʔ**áa**</u>.pa.ha. <u>ii</u>ˈtá.há.
'bird arrow' 'hurts'

17. Assuming, of course, that PrWds cannot overlap.

18. Nor, apparently *need* a second PrWd be erected, which may yield some insight into the nature of the Lx∼Pr constraint. To satisfy this constraint, it seems to be enough for even just one syllable of a PrWd to be part of the lexical item. This looks like some sort of constraint on coincidence, perhaps of the implication type that Eisner (1997) discusses, building on Cole and Kisseberth (1994):

LxWd → PrWd: ∀ LxWd *x*, ∃ PrWd *y* such that *x* coincides with *y*.

Partial coincidence is sufficient to satisfy the constraint.

19. For completeness, include undominated Main-Last: Main stress is located in the rightmost foot, mentioned in passing in Green and Kenstowicz (1995), but stated explicitly in Green (1995).

20. Kenstowicz (p.c.) suggests that Barasana pitch accent may be such a case.

References

Cole, J., & Kisseberth, C. (1994). An optimal domains theory of harmony. *Studies in the Linguistic Sciences, 24*(2).

Eisner, J. (1997). What constraints should OT allow? Rutgers Optimality Archive.

Everett, D. (1985). Syllable weight, sloppy phonemes, and channels in Pirahã discourse. In M. Niepokui et al. (Ed.), *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society* (pp. 408–416). Berkeley: UC Berkeley.

Everett, D. (1986). Pirahã. In D. Derbyshire & G. Pullum (Ed.), *Handbook of Amazonian Languages* (pp. 200–326). Berlin: Mouton.

Everett, D. (1987). Pirahã clitic doubling. *Natural Language and Linguistic Theory, 5*, 245–276.

Everett, D. (1988). On metrical constituent structure in Pirahã phonology. *NLLT, 6*, 207–246.

Green, A. (1997). The Prosodic structure of Irish, Scots Gaelic, and Manx. Ms., Cornell University.

Green, T. (1995). The stress window in Pirahã: A reanalysis of rhythm in Optimality Theory. Ms., MIT.

Green, T., & Kenstowicz, M. (1995). The Lapse constraint. Ms., MIT.

Hagstrom, P. (1995). When a Passamaquoddy unstressable schwa, that's a mora. Ms., MIT.

Hammond, M. (1987). Hungarian cola. *Phonology Yearbook, 4*, 267–269.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

Hewitt, M. (1992). Vertical maximization and metrical theory. Ms., Brandeis University.

Itô, J., & Mester, A. (1992) Weak layering and word binarity. Ms., UC Santa Cruz.

Kager, R. (1993). Alternatives to the iambic/trochaic law. *NLLT, 11*, 381–432.

Kager, R. (1995). Ternary rhythm in alignment theory. Rutgers Optimality Archive.

Kenstowicz, M. (1981). The Metrical Structure of Arabic Accent. Ms., U Illinois.

Kenstowicz, M. (1983). Parametric Variation and Accent in the Arabic Dialects. *Chicago Linguistics Society, 19*, 205–213.

Key, H. (1961). Phonotactics of Cayuvava. *International Journal of American Linguistics, 27*, 143–150.

Key, H. (1967). *Morphology of Cayuvava.* Paris: Mouton.

Levin, J. (1988). Generating ternary feet. *Texas Linguistic Forum, 29*, 97–113.

McCarthy, J., & Prince, A. (1993a). Prosodic morphology I. Rutgers University Technical Reports 3.

McCarthy, J., & Prince, A. (1993b). Generalized alignment. Rutgers University Technical Reports 7.

Prince, A. (1980). A metrical theory for Estonian quantity. *Linguistic Inquiry, 11*, 511–562.

Prince, A. (1985). Improving tree theory. *BLS, 11,* 471–490.

Prince, A., & Smolensky, P. (1993). Optimality Theory: Constraint interaction in generative grammar. Ms., Rutgers University and University of Colorado at Boulder.

Walker, R. (1996). Prominence-driven stress. Rutgers Optimality Archive.

The status of word stress in Indonesian*

Ellen van Zanten+, Rob Goedemans✓+, and Jos Pacilly+
+Leiden University / ✓Utrecht University

1. Introduction

In many cases the link from phonetic data to phonological theory is fairly straightforward. Most categorical distinctions made in the phonological systems of languages find their basis in clearly perceptible differences in the acoustic signals of the speech sounds involved. In the present article we highlight one area in which this relation has proven much less straightforward, and sometimes even highly elusive: stress. Impressionistic descriptions of this phenomenon may generally suffice, but in many cases, more thorough, experimental, research is needed to uncover the phonetic and phonological characteristics of word stress. A case in point is Indonesian. In this article we discuss some phonetic experiments carried out to resolve a long-standing phonological debate concerning Indonesian stress.

1.1 Background

Indonesian stress is phonetically only weakly marked (Teeuw 1984: 9). No phonological rules, structural or contrastive differences based on stress are observed. Moreover, it serves no contrastive functions, nor does it seem to be communicatively relevant (van Zanten & van Heuven 1998). On the other hand, deviations from the correct pronunciation "sound awkward" (Moeliono & Dardjowidjojo 1988: 73).

There is an ongoing discussion on the location of stress in Indonesian. Traditionally, most authors claim that the penultimate syllable is stressed, unless this syllable contains a schwa, in which case stress is final (Alieva, Arakin, Ogloblin, & Sirk 1991: 63; Teeuw 1984: 9). However, Laksman (1994) found evidence that schwa can be stressed as well as any other vowel. Others claim that main stress is on the final syllable of the word (Samsuri 1971) or that Indonesian has no word stress at all (Halim 1974; Zubkova 1966). According to Halim (1974: 111–113), stress location depends on the position of the word in the sentence: before a sentence-internal boundary the stress falls on the final syllable of the word preceding the boundary, whereas sentence-final stresses fall on the penultimate syllable of the last word of the sentence.

Working in a current metrical framework, Cohn (1989) and Cohn & McCarthy (1994) present a set of rules by which the patterns of main and secondary stresses in Indonesian can be derived. Their rule for main stress follows the mainstream tradition we touched upon above, while secondary stresses fall on the first syllable and every odd syllable thereafter (but never on the syllable abutting the main stress) in words of four or more syllables, as in (1).

(1)	sòlidarítas	'solidarity'	màsyarákat	'society'
	pàscasarjána	'postgraduate'	sàndiwára	'theatre, drama'

Recent investigations reveal a general preference for speakers to stress the prefinal syllable (van Zanten & van Heuven, to appear), but free variation of stress position is commonly observed, especially in longer words (van Zanten 1994: 161–163). For a more elaborate literature survey on Indonesian stress the reader is referred to Odé (1994: 39–41).

Most authors state that complex words (base plus one or more suffixes; prefixes are generally agreed not to influence the stress pattern of the base) have the stress on the penultimate syllable regardless of word-internal structure (e.g. Lapoliwa 1981: 127–131; Cohn & McCarthy 1994), but de Hollander (1984: 27–28) and Alieva et al. (1991: 64) claim that in some cases stress is maintained on the penultimate syllable of the base when a suffix is attached to it. Prentice (1994: 417) proposes a solution to this controversy that is based on the fact that Indonesian is spoken on a variety of substrate languages. This means that the pronunciation of Indonesian may differ considerably depending on the origins of the speakers. Prentice divides the Indonesian-speaking world into two regions: a 'Western' region (Kalimantan, Sumatra), where suffixation does not induce rightward stress-shift, and an 'Eastern' region (Java, Sulawesi and eastward), where stress falls on the penultimate syllable of the word, regardless of its internal structure. We follow Prentice in the assumption that the substrate language of the speaker is of crucial importance for the realisation of stress in Indonesian. For this reason, we decided to include the substrate factor in our experiments.

The combined reports on Indonesian stress sketch a confusing picture. The first aim of this study is to resolve this confusion by determining the location of Indonesian word stress through experimental research.

1.2 Prosodic considerations

Speakers can use specific variations of duration, pitch, loudness and vowel quality to highlight certain constituents within linguistic domains (Fry 1958), but usage of these prosodic properties varies across languages. The properties in question can be used to phonetically realise the abstract phonological property called stress. Any effort to track stress through its highlighting features (or *correlates*), however, is complicated considerably by the existence of another phenomenon. In languages like Dutch and English, constituents are placed in focus (i.e. marked as communicatively relevant by the speaker) by placing a salient pitch movement, an *accent*, on their prosodic head (Baart 1987; Ladd 1996). Likewise, in Indonesian, any constituent of a sentence can be put in focus by providing it with a "pitch accent" (Samsuri 1978). On the word level, the prosodic head is the stressed syllable (cf. Bolinger 1958). Hence, when a word is in focus, the stressed syllable carries the correlates of both stress and accent. Part of the confusion hinted at above may have been caused by such compounding of stress and accent correlates. So, to investigate the phonetic correlates of stress proper, we have to take into consideration the distinction between stress and accent. We must regard as true stress correlates only those cues that remain when words are not in focus (not presented as important by the speaker to the hearer). In those instances, words are not provided with a pitch accent by the speaker, and only duration, loudness and vowel quality remain as primary cues for stress.

The differences that languages show in their realisation of stress harbour yet another danger. In case the mother tongue of the researcher and the language that he investigates are different in this respect, we advocate caution, because the perception of the linguist may be coloured by the stress rules of his own language. In this light, it seems appropriate to perceptually test the native-speakers' intuition on stress position. Such intuitions have to result from carefully controlled perception experiments in which judgements can be indirectly obtained, even in cases where subjects would normally be unable to indicate stresses, or where stresses are not acoustically marked (Berinstein 1979). In this paper we present two experiments in which we obtain such indirect native-speaker judgements from Indonesian listeners for a variety of stress patterns.

In the first experiment, Indonesian listeners with different regional backgrounds are asked to indicate which of two stress patterns they prefer. The second experiment is an evaluation test in which the same listeners are asked to rate the acceptability of different stress patterns. If Indonesian words are preferred and judged more acceptable when the (main) stress is realised on the prefinal syllable we will conclude that the traditional rule (stress is fixed on the prefinal syllable) is correct. If, on the other hand, acceptability and preference are not significantly influenced by stress position, word stress must essentially be free. Whatever the outcome may be, if the results of our experiments can resolve the controversy concerning Indonesian word stress, we will have demonstrated the necessity of phonetic experimentation as the foundation of phonological theory, which is the second aim of this study.

In the next section we discuss the production experiment we carried out to obtain the necessary data on Indonesian stress correlates for the two perception experiments described here.

2. Production data

2.1 Method

To investigate stress position and its relevance in Indonesian, we used resynthesized speech in which the correlates of Indonesian stress were manipulated. To determine which correlates should be used, and in which proportions, a speech-production test was devised, for which the words in (2) were selected.[1]

(2) 4σ	masyarakat	'society'	kacamata	'spectacles'
	laksamana	'admiral'	dikatakan	'it was said'
	perempuan	'woman'	cendekia	'clever'
5σ	pascasarjana	'postgraduate'	solidaritas	'solidarity'

The target words were embedded in the carrier sentence *Dia mengucapkan kata (solidaritas)*, 'He pronounces the word (*solidaritas*)'. Target words are thus in sentence final (focus) position and expected to receive an accent-lending pitch movement on the stressed syllable (van Heuven 1994: 15; Samsuri 1978). Secondly, to get an idea of the correlates of stress proper, the targets were embedded in non-final (non-focus) position in the carrier sentence: *Kata (solidaritas) itu tepat*, 'The word (*solidaritas*) is correct'. To keep the proliferation of stimulus material in check, we did not include the sentence final non-focus condition, nor the sentence medial focus condition.

The target words in their carrier sentences were each read twice by two male Indonesian speakers. One of the speakers had a Javanese background; he originated from Klaten (Central Java), and had come to the Netherlands quite recently. *Javanese* is considered to be the most influential regional language of Indonesia (Poedjosoedarmo 1982; Steinhauer 1980). In 1990, over one-third of the Indonesian population spoke Javanese as a first language (Steinhauer 1994: 781). As some of the sources claim for Indonesian, the penultimate syllable is weakly stressed in Javanese, unless this syllable contains a schwa, in which case stress is shifted to the final syllable (Ras 1982: 13). Poedjosoedarmo, on the other hand, seems to hold the view that stress is final in Javanese (Poedjosoedarmo 1982: 49; Note 45). Our second speaker was a Toba Batak who had lived in the Netherlands for some years but spoke Toba Batak as well as Indonesian frequently. *Toba Batak* differs crucially from Javanese in that stress can be contrastive, and in that the stressed syllable – usually the penult – is clearly marked by prosodic means (Nababan 1981: 27, 135).

All material was recorded on DAT with a Sennheiser MKH 416 unidirectional condenser microphone and transferred to a Silicon Graphics workstation (downsampled to 16 kHz), and stylized and resynthesized ('t Hart, Collier, & Cohen 1990), after which the relevant pitch movements and segmental durations of the target words were measured. Peak intensity was measured for all syllables in the targets. The next section contains a summary of the results found in the production experiment, to the extent that these are relevant for the perception experiments.

2.2 Results and discussion

The production data of our Toba Batak speaker showed a clear influence of this speaker's substrate language. In each target word, one syllable, predominantly the prefinal, was significantly longer and louder – both in and out of focus – than the other syllables. When in focus, vowels in such prominent syllables were spoken on a higher pitch. Pitch can only be the auditory correlate of accent if the pitch movement is steep and minimally 3 semitones (a semitone (st) is a 6% difference between two frequencies) and occurs in a specific position within the (stressed) syllable (van Heuven 1994: 19). This is exactly what we found for our Toba Batak speaker: the pitch movement consisted of a steep rise of approximately 3.5 st which started at the end of the preceding vowel. The high pitch plateau lasted for the full length of the accented vowel.[2] It was followed by a 9 st pitch fall. Moreover, the average duration of vowels in accented syllables was approximately 50% longer than the average duration of unaccented vowels. For consonants, lengthening was around 25%. Such lengthening effects were also attested for Toba Batak speakers by van Zanten and van Heuven (1997: 210–211); they fit in well with data on (stress) languages like English and Dutch; cf. Nooteboom (1972); Eefting (1991). Finally, peak intensity in the accented syllable was 2.5 dB (decibel) higher, on average, than in the unaccented vowels in the Toba Batak speech data. It was usually the penultimate syllable that was made prominent in this way. A notable exception was the loan word *solidaritas* 'solidarity', in which the last syllable stood out.

The speech data for the Javanese speaker were quite different. Acoustically, no duration or intensity differences between syllables were found that could be related to a stress pattern. As regards pitch, in the [+focus] condition, a small pitch rise (around 2 st) was usually found on the first syllable of the target, and a pitch fall (of approximately 8 st) somewhere near the boundary between the prefinal and the final syllable (cf. also Ebing 1997); this apparently common Indonesian pitch contour is also reported in van Zanten and van Heuven (to appear). Pitch rise nor fall meet the requirements for accent perception in stress languages. The rise is well below the threshold excursion size of around 3 st, and the fall is not in a specific position in the syllable. However, impressionistically, there was a tendency for the fall to lend prominence to the final or penultimate syllable, depending on its starting point. At this stage we kept open the possibility that this is the way in which Javanese speakers realise word stress.

The speech data we collected from the two speakers thus differ fundamentally with respect to stress realisation. This can be related to the background languages of the speakers, viz. Toba Batak, a (lexical) stress language, and Javanese, a language for which stress is described as weak, and for which there is some debate concerning its location (see Section 2.1). For the Toba Batak speaker we measured steep and sharply defined pitch movements as well as duration and loudness effects which could be related to prominent syllables. For the Javanese speaker, on the other hand, no such systematic effects were found.

3. Perception experiments

3.1 Stimuli

The perception tests were carried out with the six four-syllable, and two five-syllable target words in (2). These target words were manipulated in accordance with our findings in the production data. We constructed two sets of stimuli, one based on the production of the target words by the Toba Batak speaker, the other on the utterances of the Javanese speaker, spoken in sentence-final (focus) position.

In order to investigate the acceptability of variable stress positions in Indonesian words we wished to compare the judgements of native speakers on stimuli with a prominent prefinal syllable (i.e. stressed according to the traditional rule; e.g. *masyarákat*) with stimuli in which one of the other syllables in the same word was made prominent (*másyarakat, masyárakat* and *masyarakát*). The number of stimuli for each target word thus depended on the number of syllables in that word. In addition, we devised for each target word one stimulus in which none of the syllables was made prominent, but in which the first syllable of the preceding word *káta* was accented. We expected this stimulus to score low in the tests. Altogether, five stimuli for each of the four-syllable words were created, and six stimuli for the five-syllable words. These included a "0" stimulus in which the first syllable of *káta* was made prominent in the same way, but none of the syllables of the target.

For the Toba Batak speaker prominence was achieved by manipulating pitch, duration and intensity in accordance with the mean values which we found in the Toba Batak speech data. First, prominence differences in each target word were removed. For each stimulus, one syllable was then made prominent in the following way. Vowels were lengthened by 50%, consonants by 25%, and intensity was raised by 3 dB. Pitch movements were adapted as indicated in the top panel in Figure 1, which schematically represents the pitch contour for the Toba Batak stimuli with prominent penultimate syllables.

Similarly, the Javanese stimuli were based on the Javanese speech data. There was no variation in durational structure or intensity that could be attributed to stress position for the Javanese speaker. The (start of the) fall was the only possible prominence lending cue. So, we decided to vary its position, leaving the inconspicuous rise in a fixed position. Thus, the pitch contour consisted of a 2 st rise on the first syllable of the target word, followed by a high pitch plateau of variable duration, and an 8 st fall. These manipulations of the pitch contour are schematically represented in Figure 1 (bottom panel). In the Javanese speech data, the pitch fall often occurred somewhere in the border region between penult and final syllables. To enable subjects to choose in a clear-cut way between accent positions, we devised the stimuli such that the fall started within one specific syllable each time. A "0" stimulus was created here as well, for which the first syllable of *káta* was made prominent. Finally, we included an exact (stylized) copy "S" of the original pronunciation of each of the eight target words in which the fall occurred in the border region between the prefinal and final syllables.

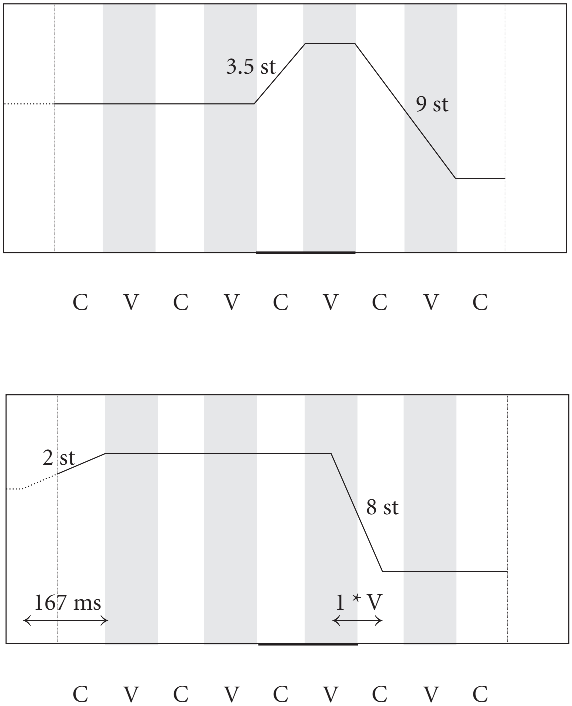


Figure 1. Schematic representation of the pitch contours for the Toba Batak (top panel) and Javanese (bottom panel) stimuli with prominent penultimate syllables (indicated by the bold line on the x-axis).

All stimuli were superimposed on the mean pitch declination of the original utterances: a downtrend of 1 st per second for both speakers. To reduce the workload on the listeners, the first part of the carrier sentences (*Dia mengucapkan*) was deleted.

3.2 Experimental paradigms

Two types of listening experiments were devised. The first was a pairwise-comparison experiment in which subjects are requested to choose between two members of a stimulus pair on the basis of a parameter that has been varied in these stimuli. In our case, the pairs consisted of a reference stimulus (with prominent prefinal syllable) and any one version of the same target word. For the Toba Batak-based speech this amounted to 84 pairs of stimuli (that is, $6 \times 5 + 2 \times 6 = 42$ stimulus pairs in both orders). The Javanese-based stimuli consisted of 100 stimulus pairs (42 + 8 stylized versions "S", in both orders). Stimuli were recorded on audio tape in quasi-random order. Secondly, we devised an evaluation test in which subjects were to judge the acceptability of the individual stimulus words. Each individual stimulus was copied twice on tape in counter-balanced random orders. Consequently, the number of judgements asked was equal to that of the previous experiment: 84 for the Toba Batak-based stimuli, and 100 for the Javanese-based stimuli.

3.3 Listeners

Three groups of Indonesians took part in the listening experiments. Two groups were selected to match the substrate languages of the original speakers, i.e. a group of 20 speakers of Indonesian who had Javanese as a substrate language, and a group of 13 Indonesian speakers with a Toba Batak substrate. We decided to add a third group, which consisted of 9 speakers of Indonesian who lived in Jakarta and spoke colloquial (Jakartan) Indonesian or Jakarta Malay (Betawi, Jakartanese) but no other regional language. As the language of the capital which is regularly heard in national radio and television programs, Indonesian as spoken in Jakarta is thought to have an enormous influence on (Standard) Indonesian (Steinhauer 1980; Adelaar & Prentice 1996: 678). Within Jakarta, various dialects are spoken (cf. Grijns 1991; Ikranagara 1980; Muhadjir 1981). However, we found no evidence in the literature that these dialects differ with respect to stress placement. Wallace (1976) specifically studied phonological variation in Jakarta Malay. According to him, accent is a property of phrases rather than of words in Jakarta Malay. In the last word of a phrase, accent is placed by a rule "common to many Malay dialects: if the penultimate syllable contains schwa, accent the syllabic element of the last syllable; otherwise, accent the syllabic element of the penult"[3] (1976: 58). Wallace makes no mention of suprasegmental variation between Jakartan speech varieties. In selecting listeners, we therefore did not distinguish between, for instance, subjects who said that they spoke Jakarta Malay or Jakartan Indonesian. It seemed sufficient to exclude subjects with non-Jakartanese regional substrates from our third group of listeners.

3.4 Procedure

The tape was played to the listeners over good-quality earphones at a comfortable listening level. Eleven listeners were tested individually at the Phonetics Laboratory of Leiden University and the others were tested in two groups in a language lab in Jakarta. Subjects were told that they were going to listen to the final parts of declarative sentences and that these had different intonations (*lagu kalimat*). They were not informed about the actual purpose of the experiments, i.e. to compare the acceptability of different stress patterns. For the paired comparison test, listeners were asked to indicate on their answer sheets which of each pair they preferred. It was made clear to them that they had to make a single choice in all cases; blanks were not allowed. For the acceptability test, subjects were instructed to rate the acceptability of each phrase on a ten point scale, ranging from 1 ("very bad") to 10 ("very good"). They were requested to circle the appropriate mark on the answer sheets for each stimulus phrase. Each test was preceded by three practice items. After this the tape was stopped to answer any questions raised by the listeners. After every ten items a short beep was inserted to help the listeners keep track of the stimuli on their answer sheets. All instructions were in Indonesian. Approximately half of the subjects (i.e. half of each listener group) listened to the "Javanese" stimuli first (first the pairwise-comparison test and then the acceptability test) and then to the "Toba Batak" pairwise-comparison and acceptability tests, respectively. The other subjects were presented with the "Toba Batak" stimuli before listening to the "Javanese" stimuli; again, the pairwise-comparison test was followed by the acceptability test.

4. Results and discussion

4.1 Javanese-based stimuli

4.1.1 *Pairwise-comparison experiment*

In the pairwise-comparison experiment, each stimulus pair consisted of two instances of the same word; one of the comparison stimuli and the reference stimulus with the penultimate stress. For each substrate-listener group, and all possible stress positions, we calculated the percentage of responses in which the subjects chose the comparison stimulus as the better one of the pair. In case the reference and comparison stimuli are identical, the outcome should be exactly 50%, because it is impossible to select one stimulus as better than the other in that case. In practice, however, subjects tend to choose the first member of the pair when they are unable to make a motivated choice. This bias for the left-hand member of a stimulus pair is known as the Time Order Error (cf. Woodrow 1951; van Heuven & van de Broecke 1982). In our experiments we eliminated the TOE effect by presenting the stimuli to the subjects in both orders.[4] If Indonesian does indeed have penultimate stress, all the percentages representing responses for non-identical stimulus pairs should lie well below this 50% mark. In these cases, the comparison stimulus does not have the preferred penultimate stress, and should, therefore, not be selected as the better one of the pair.

In this section, and the next, we concentrate on the stimulus words with four syllables, disregarding any differences between the individual words. The results for the three groups of listeners and the Javanese-based stimuli are presented in Figure 2. In this figure, the percentage score for the comparison stimulus (Scomp) is plotted along the y-axis, and the stress locations are placed on the x-axis (remember that in the "0" case, stress is located on the first syllable of *káta*, "S" represents the stylized version). The scores for each of the three substrate groups are connected by lines.

We observe that the data points for the *Javanese* listeners do not follow the pattern for penultimate stress outlined above. Stress on the final syllable is judged as acceptable as stress on the penultimate syllable. A one-way analysis of variance (anova, $\alpha = .05$) shows that there is a significant difference in the percentage scores; $F(5,1290) = 37.97$, $p < .001$. A post-hoc SNK analysis shows that this is attributable to a difference between a "0", "1" and "2" group on the one hand, and a "3", "4" and "S" group on the other. Assuming, for the mo-

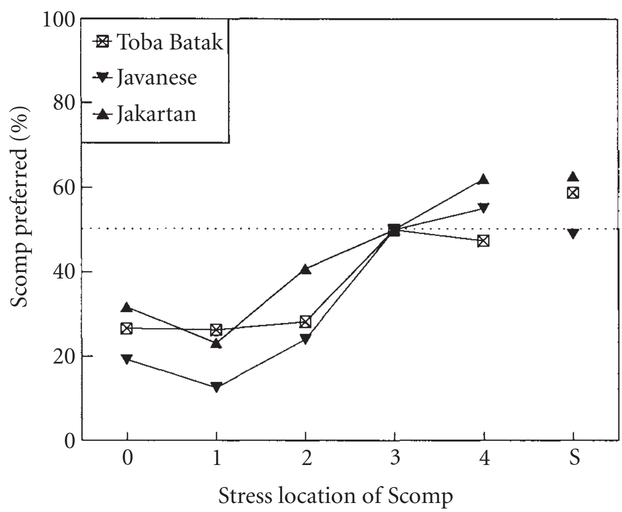


Figure 2. Javanese-based 4-syllable stimuli. Percentage of cases in which the comparison stimulus (Scomp) was judged better than the reference stimulus. Broken down by stress location and substrate listener group.

ment, that we are still dealing with stress patterns in Indonesian, we interpret this as the difference between acceptable and unacceptable stress locations. We postpone the discussion of the status of these stresses to Section 5.1.

We find the same pattern for the other two substrate groups. Both final and prefinal stress are acceptable, but stress on the first and second syllable (or no stress at all) is not. One-way anovas again reveal significant differences in the percentage scores: $F(5,642) = 14.76$, $p < .001$ for the *Jakartans* and $F(5,930) = 20.83$, $p < .001$ for the *Toba Batak* listeners. The post-hoc SNK analysis again distinguishes two groups in both cases.

Within the "unacceptable" group, "2" is judged better than "1" by both the Javanese and the Jakartan listeners. We do not yet know whether this tendency is linguistically relevant. Within the "acceptable" group, both the Javanese and the Jakartan listeners, as opposed to the Toba Bataks, seem to prefer stress on the final syllable over stress on the penultimate syllable. This tendency goes against the claim that stress is predominantly penultimate in Indonesian. Finally, the three groups of listeners all judge the stimuli with penultimate or final stress to be approximately as acceptable as our close-copy version of the Javanese pitch contour. We take this as an indication that these manipulated stimuli sounded sufficiently natural to Indonesian ears.

So far, the evidence we have found argues against the claim that stress is penultimate in Indonesian. Let us now look at the data from our evaluation experiment to see whether these point in the same direction.

4.1.2 *Evaluation experiment*

Figure 3 shows the mean evaluation scores, represented on a scale from 0 to 10 on the y-axis, for all stress locations (x-axis), again broken down by substrate group, as in Figure 2.

The lines connecting the scores for each group largely follow the same pattern as those in Figure 2. Stress locations "3", "4" and "S" again form one group which is significantly different from the combination "0", "1" and "2" (anova: Javanese $F(5,1577) = 106.1$, $p < .001$; Toba Batak $F(5,642) = 28.3$, $p < .001$; Jakartan $F(5,642) = 26.6$, $p < .001$; post-hoc SNK for all groups). This figure shows that the difference between the stress location groups is more pronounced for the *Jakartan* listeners. The transition between the groups is gradual for the *Javanese* and the *Toba Batak* listeners. This trend is confirmed in the post-hoc statistical analyses.

The results from our evaluation experiment corroborate the findings of the pairwise-comparison experiment. We continue with an analysis of the five-

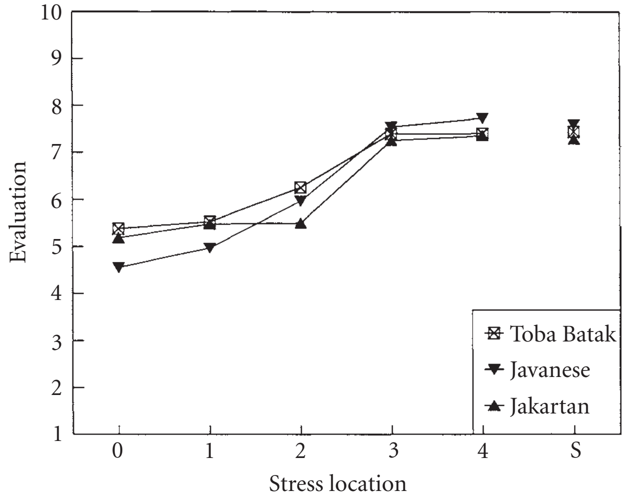


Figure 3. Javanese-based 4-syllable stimuli. Evaluation scores for all stress locations, broken down by substrate listener group.

syllable words, to see whether the results are compatible with those found for the four-syllable words.

4.1.3 *Five-syllable words*

Originally, the five-syllable words were incorporated into our experiments to check how far the main stress could be moved to the left edge of the word, while remaining acceptable to Indonesian listeners. We have already seen that stress on syllables too far from the right-hand word edge is judged to be only marginally acceptable in words of four syllables. We expect this judgement to be the same for five-syllable words.

The results we found for the Javanese-based stimuli of five syllables confirm our expectations. Although there are only two words in this subset, the general pattern is obvious for all three substrate groups of listeners. In Figure 4, the results for the pairwise-comparison experiment and the evaluation experiment are presented.

For all groups of listeners the scores for stress locations "0", "1" and "2" were significantly lower than the score for the locations "4", "5" and "S", in both experiments (one-way anova with post-hoc analysis). A recurring tendency we found with respect to stress location "3" was that it seemed to be judged as transitional between the two groups above. Its stressability possibly depended on the target word. On the whole, however, our stress manipulations in the two target words were perceived in the same vein: stress on the final or the penultimate syllable is preferred over stress on any one of the other syllables.[5] The stylized version, for which our "Dutch ears" cannot determine whether stress is final or penultimate, is also considered to be acceptable. In short, the results

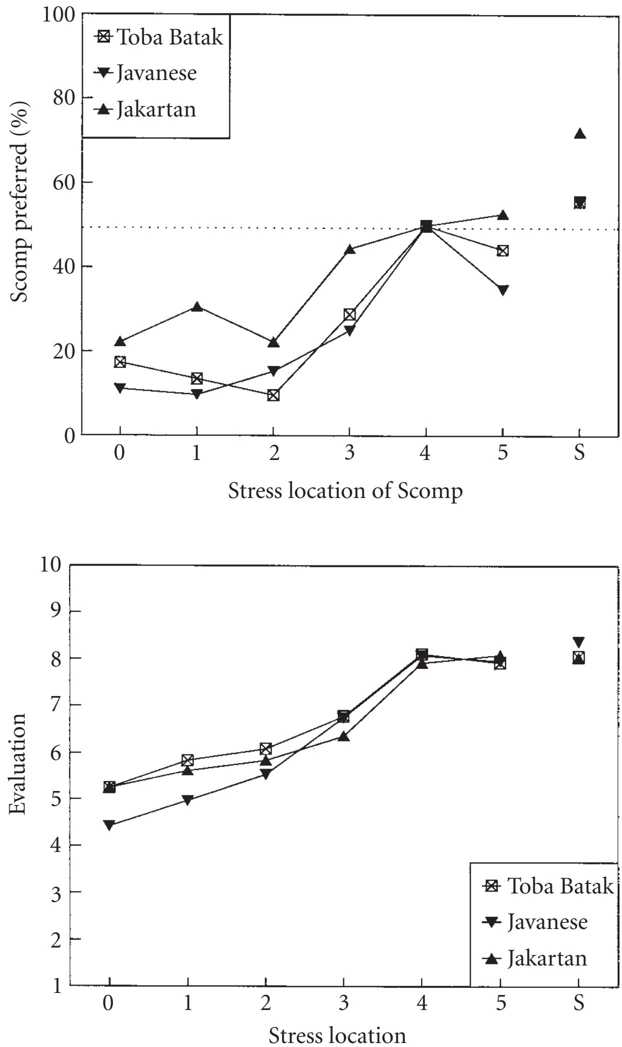


Figure 4. Pairwise comparison (top panel) and evaluation (bottom panel) results for Javanese-based stimuli of five syllables, each broken down by stress location and substrate listener group.

for the Javanese based five-syllable words confirm the conclusion we drew from the four-syllable words: stress in Indonesian, as spoken by the Javanese speaker, is acceptable on either the final or the penultimate syllable in all the words we included in our experiments. Such variation suggests that Indonesian has free stress, which is tantamount to having no stress at all (van Heuven 1994: 18). Therefore, we suspect that some phenomenon *other* than stress is responsible for the facts we observe here. Remember, in this respect, that we manipulated only those properties of prominent syllables in focus that we had found to be relevant in the production experiment. For the Javanese speaker, this boiled down to a manipulation of pitch only. Let us consider the possibility that such manipulations were not perceived as stress manipulations. If that is the case, we must also consider the subjects' reactions to a form of Indonesian that does have "real" stresses, in order to be more conclusive about the perception of stress by the three substrate groups. We have determined in the production experiment, that the Toba Batak speaker realises stress in [+focus] condition with a pitch movement, a rise in intensity and lengthening. Considering the speaker's substrate language (which has lexical stress), and his phonetic realisation of the word stresses, we are positive that the Toba-based stimuli contain the desired differences in stress location. Let us see how the subjects reacted to these stimuli.

4.2 Toba Batak-based stimuli

4.2.1 *Pairwise-comparison experiment*

As for the Javanese-based stimuli we start with the analysis of the words with four syllables. Figure 5 presents the pairwise-comparison data for the Toba Batak speaker in the fashion of Figure 2. Let us first consider the *Toba Batak* listeners. The figure shows that they clearly prefer stress to fall on the penultimate syllable. The percentage score for stress on the penultimate syllable, which is slightly, though not crucially, flattered by the bias, is shown to be different from all the other scores in a one-way anova: $F(4,775) = 14.43$, $p < .001$ with post-hoc SNK analysis.

We interpret this result as a reflection of the Toba Batak stress rule in the Indonesian spoken by the Toba speaker. It seems that the Toba listeners prefer penultimate stress in this case. We will come back to this issue in the general discussion.

The *Javanese* listeners react crucially different to the Toba Batak stimuli. The percentage scores for all stress positions are much closer together than for the Toba listeners. Only the score for stress on the second syllable is significantly

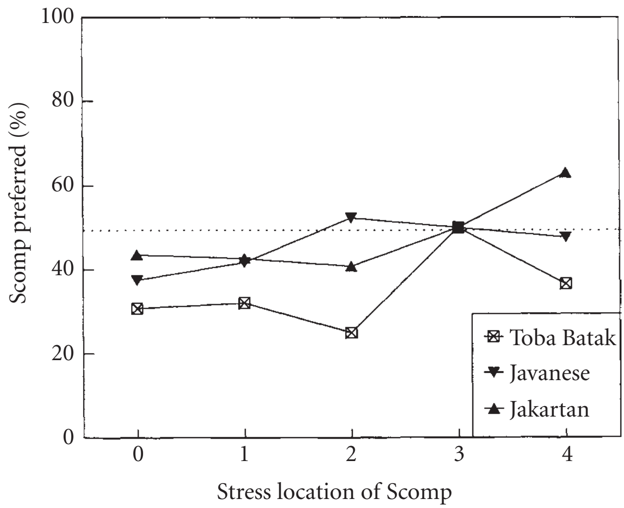


Figure 5. Toba Batak-based 4-syllable stimuli. Percentage of cases in which the comparison stimulus (Scomp) was judged better than the reference stimulus. Broken down by stress location and substrate listener group.

different from the score for stress on *káta*. There is no significant difference found between any of the scores for stress *within* the target word. Apparently, a Toba-style stress realisation is equally acceptable to Javanese ears on any syllable of a four-syllable word.

Finally, the *Jakartan* listeners judge the Toba Batak-based stimuli in much the same way as the Javanese listeners. However, in this case, stress on the final syllable is statistically different from all the other cases. Let us now see whether the evaluation data corroborate the pairwise-comparison data for the Toba Batak-based stimuli, as they did for the Javanese-based stimuli.

4.2.2 *Evaluation experiment*

The evaluation scores for the Toba Batak-based stimuli by our three groups of listeners are shown in Figure 6. The preference for the prefinal syllable by the *Toba Batak* listeners, which we found in the pairwise-comparison test is confirmed by the high score for this syllable (7.5). Statistically however, "3" differs only from "0" and "4".

The evaluations of the *Javanese* listeners resemble the results for the Javanese listeners found in the pairwise-comparison experiment quite closely. The rating for stress on the second syllable is again somewhat higher than the others, but this difference is not significant. The only significant difference is

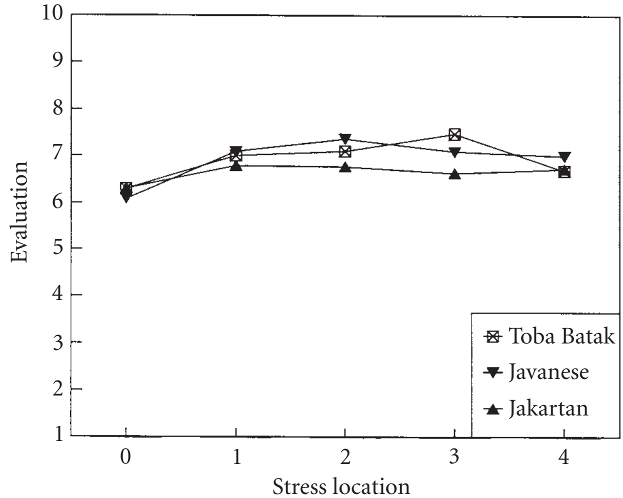


Figure 6. Evaluation scores for the Toba Batak-based 4-syllable stimuli for all stress locations broken down by substrate listener group.

that between all the possible stresses on the target word on the one hand, and stress on *káta* on the other, indicating that, for Javanese listeners, Toba Batak style stresses are equally acceptable on all syllables in the word.

The slight preference for stress on the final syllable, which we found for the *Jakartan* listeners in the pairwise-comparison experiment is not reflected here. There is no significant difference between any of the evaluation scores for the Jakartan listeners.

As in the pairwise-comparison experiment, the Javanese and the Jakartan listeners seem indifferent to where the stress falls in Indonesian four-syllable words pronounced by the Toba Batak speaker. We suppose that the difference in stress realisation between the Javanese and the Toba Batak speaker is responsible for this difference in perception, but before we elaborate on this difference, we will briefly discuss the five-syllable words of the Toba Batak speaker.

4.2.3 *Five-syllable words*

Turning our attention to the five-syllable words, we did not find any significant groupings for the two stimulus words together. This seems to be caused by the fact that these words may have different preferred stress patterns, i.e. final stress for *solidaritas* and initial or penultimate stress for *pascasarjana*. We believe that this difference is caused by a crucial difference in status of these two words; the former is a recent loan, and the latter is morphologically complex (*pasca+sarjana* 'post+graduate'). Therefore, we should look at these words separately in our analysis. Unfortunately, too few responses for the separate words were collected to justify statistical analyses like the ones given for the other cases. No firm conclusions for the Toba Batak-based stimuli with five syllables can be drawn at this point.

5. Discussion

5.1 The status of word stress

Recapitulating the previous sections, we briefly consider the importance of the results presented. Firstly, as Figures 2 and 3 clearly show, listeners from three different backgrounds judge the "Javanese-style" stress on either the final or the penultimate syllable to be acceptable in words of four syllables. This is confirmed for words of five syllables in Figure 4. Finally, Figures 5 and 6 show that the Jakartan and Javanese listeners are indifferent to the location of "Toba Batak-style" stresses. The Toba Batak listeners themselves clearly prefer penultimate stress in this case.

Thus, we observe three different reactions to our two types of stimuli. As we expected, these reactions depend on the substrate language of the listener group, as well as the difference between the stimulus types. Let us consider the difference in stimuli first. As was noted above, in the stimuli that were based on the speech of the Javanese speaker, stress in focus was realised as a pitch movement only. The stimuli based on the Toba Batak speaker contained more canonical stresses, realised by greater intensity and longer duration as well as pitch movement. Since these differences in the stimuli elicited different reactions, we hinted at the possibility that the parameter we varied in the Javanese stimulus set represented something other than stress. The difference in realisation of the two stress types suggests that we must look upon the Toba Batak stress as the "real" one. The reactions to the two types of stimuli confirm this. In the Javanese-based stimuli, the "stress" may fall on either the penultimate or the final syllable in the same word, which indicates that we are dealing with a relatively free stress, or, to push the point even further, no stress at all. For the moment, let us label the variable parameter in these stimuli with the cover-term *prominence* instead of stress.

Turning our attention, then, to the influence of substrate listener group on stress perception, we may try to explain some of the differences we observe while keeping in mind the split between prominence and stress that we introduced above. Stress locations other than penultimate in the Toba Batak-based stimuli are correctly rejected by the Toba Batak listeners themselves because they are used to hearing such stresses and expect them to be mostly penultimate in their substrate language. The Javanese and Jakartan listeners, however, do not differentiate between the different stress locations. So, if the set of acoustic syllabic properties that we generally define as stress appears, they either cannot hear, or do not care, on which syllable these properties are realised. We can simply say that stress has no meaning to them, which is all the more reason to assume that word stress is neither a feature of Javanese and Jakartan, nor of the variant of Indonesian they speak.

When we consider their reactions to the prominence differences in the Javanese-based stimuli, we find that the Javanese and the Jakartan listeners are not totally insensitive to intonation movements. Instead, they wish the prominent intonation movement to occur at the right edge of the word and judge stimuli to sound gradually worse as prominence moves further to the left edge of the word. Remarkable, in this respect, is the behaviour of the Toba Batak listeners. We expected them to interpret the Javanese intonation movement as a stress, and consequently, allow it to occur only on the penultimate syllable. The fact that they copy the behaviour of the Jakartan and Javanese listeners in this case, and do accept final prominence, indicates to us that they enter a different "mode" when listening to the group of speakers that defines (through group size, but also through great influence in politics and the media) the most common form of Indonesian. We should mention here that all but one of the Toba Batak listeners lived in Jakarta.

We note that, with respect to stress (or prominence) there is no uniform rule for Indonesian. A long history of debates on the exact location of stress is indicative of the absence of such uniformity. We add to this evidence the large differences in the reactions of listeners with different substrate backgrounds to variants of Indonesian stress produced by speakers with other substrate backgrounds. However, if we are forced to choose one particular stress or prominence rule for Indonesian, it would be the rule used by the influential Javanese and Jakartans. This choice is motivated by the fact that the Toba Batak listeners react to the Javanese variant of Indonesian in the same way the Javanese themselves do.

In conclusion, we state that neither the indifference of Javanese and Jakartan listeners to word stress location, nor the acceptance of penultimate and final prominence by all three groups of listeners, point in the direction of a strictly penultimate stress. If anything, the data we have gathered in this experiment forcefully refute the claim that stress is predominantly penultimate in Indonesian. We did, however, find it to be (almost) strictly penultimate in one of its variants, in which Toba Batak listeners were subjected to the Indonesian of a Toba Batak speaker, but we can hardly claim that the pattern of this variant can be generalised over the entire Indonesian community. As we have shown, there is no "Indonesian" with respect to stress; other substrate groups might speak Indonesian with other stress patterns yet again. However, even if we took the term "Indonesian" to cover its most common variant, the claim that stress is penultimate cannot be maintained. On the contrary, we have found compelling evidence for the claim that this variant does not have stress at all.

5.2 Phrasal accent

In the previous section we stated that the phenomenon underlying the Javanese prominence pattern is not word stress. We find strong support for this claim in the fact that prominence may freely occur on any of the last two syllables (and sometimes even the antepenult) of any word, an observation that was made earlier by van Zanten and van Heuven (to appear), albeit in experiments on a somewhat smaller scale. We cannot conclude this article without presenting at least a speculative account of what we take the Javanese prominence pattern to represent. We propose that it represents a phrase-based accent.[6] Recall from Section 1 that we reserve the term accent for the instrument of focus on the phrasal level. This accent usually docks on the syllable that is stressed by abstract phonological rules or lexical specification. In the Indonesian case, however, there is no such stressed syllable. The accent may freely dock anywhere at the right edge of the phrase.

There are several signals that force upon us the idea that Indonesian (in its Javanese variant) may have a phrasal accent only. We count the evidence against Indonesian prominence as a word-level phenomenon, which we presented above, among these signals, but we will not present it again here. So, firstly, the only cue for accent is a pitch movement (see Section 1.2). In the prominent syllables of the Javanese speaker, we have found no cue other than pitch movement, reducing the likelihood of their being stressed on the word level, but keeping open the possibility that they are accented on the phrase level.

Secondly, and more importantly, the acceptability of prominence tends to rise as it occurs further towards the right edge of the word. This phenomenon is clearly indicative of a phrasal accent that should occur near the end of the phrase. An important observation in this respect is that the location of the accent is not even bound to the syllable positions. In our production experiment we found many cases in which the pitch fall occurred in between the penultimate and the final syllable. Some of these cases formed the basis for the stylized versions ("S") of the words we included in the experiment alongside the stimuli. These stylized versions were invariably judged as good as the stimuli in which the pitch fall was clearly located on either the penultimate or the final syllable.

Thirdly, the relative importance of the first syllable in five-syllable words that we found both in the production and the perception experiment (and which Cohn 1989 also reports) reminds us of the *hammock* pattern that is a feature of phrasal stress. In this "eurhythmical" pattern "stresses are spaced not too closely and not too far apart" (Hayes 1995: 372). Such a pattern is often achieved by reordering stresses that were assigned at the word level to the extreme edges of the words.[7] Our assumption of phrasal accent harbours an explanation for some other exceptional patterns we observed above as well. The Indonesian phrasal accent is relatively unrestricted, and may freely occur on any syllable other than the final or penultimate without incurring a violation of the rule. Early accents would be less acceptable, as we observe in the figures representing the general pattern, but not altogether wrong.

Finally, the phrasal accent hypothesis straightforwardly explains the behaviour of our listener groups. The Toba Batak listeners seemed to be in a different "mode" when listening to the Javanese-based stimuli. In that mode, they were apparently judging phrasal accents, a task that should not be beyond them, since they are well acquainted with the Javanese-style prominence patterns and they have proven to be quite sensitive to differences in pitch movements (van Heuven & van Zanten 1997). The Javanese and Jakartan listeners, on the other hand, cannot judge the Toba Batak word-level stresses because they are used to hearing phrase-level prominence patterns only (they considered the stresses "too harsh").

On the basis of our experimental results we claim that word-level stress in two of the most important varieties of Indonesian is totally absent. According to us, a phrasal accent is the most likely alternative. We have shown that our results are compatible with this alternative.

6. Conclusion

The most important conclusion we draw from the results of our experiments is that there is no reason whatsoever to assume that stress in Indonesian always falls on the penultimate syllable if it contains a full vowel. We have shown that speakers with different substrate languages behave differently with respect to stress realisation and perception. Even if we set this caveat aside, however, and concentrate on the variety spoken by the most dominant substrate groups (Javanese and Jakartans), we conclude that there is good reason to exclude the penultimate stress hypothesis. In our view, the rule that drives prominence patterns in the influential Javanese variety of Indonesian is phrasal. Possibly the only phonological rule that is relevant for accent location in Indonesian states that it must occur somewhere at the right edge of the phrase. Halim's (1974) claim that stress location depends on sentence position was already suggestive of such a phrasal phenomenon.

As we have seen in the introduction, some alternative stress rules have also been proposed for Indonesian. We believe that this variation is caused by the differences in prominence patterns that we observe in speakers with different substrate languages, but in some cases also by the fact that the linguists in question have tried to construct stress rules for something that is not stress at all. In general, we are wary of phonological analyses that are based on direct interrogation of (a few) native speakers, or even only the perception of the linguist. In the Indonesian case, a question like "Is sòlidarítas pronounced right?" may have been answered positively, but not because the word was uttered with the only possible stress pattern. The impression to the non-native linguist may have been that there should be a stress on the penultimate syllable, but the native listeners appear to perceive things quite differently. This research constitutes one example of the crucial importance of careful phonetic experimentation in such cases. It provides a basis for phonological claims, and can be instrumental in the resolution of long-standing phonological debates.

Notes

* We would like to thank the staff of the Erasmus House in Jakarta, in particular Herman Poelman, as well as Wim van Zanten, for running the perception experiments. Moreover, we thank Vincent van Heuven for his constructive comments.

1. Actually, we used 20 words in the production experiment. The other 12 words were included to check phonological and morphological claims that will not be discussed in this article.

2. An exception was the accented final syllable, in which the rise was immediately followed by a fall. Only in this way can the low pitch indicating the end of the sentence be audible.

3. "Impressionistically, this accent seems to consist primarily of a rise in pitch, often accompanied by an increase in duration and intensity." "Some foreign words have irregular accent: *bioskóp* 'cinema', *Amérika* 'America'" (Wallace 1976: 58–59).

4. Eliminating the TOE effect for the *identical* reference stimulus pairs proved to be impossible. We opted instead to calculate the percentage of choices for the left-hand member of the pair to obtain a measure for the size of the TOE effect in our experiment, which proved to be 9.5%. In the figures below, the score is set at the theoretically motivated 50%, which is also used in the interpretation of the statistics.

5. We cannot explain the relatively low score that the Javanese listeners produced for stress on the fifth syllable. Such a low score for the final syllable is not found in the evaluation experiment, nor in the results for the four-syllable words. It might result from the unacceptability of stress on the final syllable of *pascasarjana* (for the Javanese listeners only). Similarly, the relatively high score for stress on the first syllable for the Jakartan listeners may indicate the possibility of a stress on the first syllable of *pascasarjana*. At this stage, we cannot exclude the possibility that the patterns for individual words can deviate from the general pattern we found.

6. In the following discussion we use the term **phrasal accent**. We note here that we use this term rather loosely, denoting an intonation movement that signifies the end of a phrase. We do not wish to commit ourselves to any of the definitions for phrasal accent that can be found in the literature before we have investigated the properties of the Indonesian phrase-based accent in greater detail. In Goedemans and van Zanten (to appear) we delve deeper into this matter, and investigate the relation of the data presented here to Indonesian accent in general.

7. Note that our Javanese speaker produced small F_0-rises on the first syllable as part of the intonation pattern, which are clearly important in this respect. We used these rises in the construction of our stimuli.

References

Adelaar, K. A., & Prentice, J. (1996). Malay: its history, role and spread. In S. A. Wurm, P. Mühlhäusler, & D. T. Tryon (Eds.), *Atlas of languages of intercultural communication in the Pacific, Asia and the Americas*, Vol. II.1 (pp. 673–693).

Alieva, N. F., Arakin, V. D., Ogloblin, A. K., & Sirk, Yu. H. (1991). *Bahasa Indonesia: Descripsi dan Teori.* Yogyakarta: Penerbit Kanisius.

Baart, J. L. G. (1987). Focus, Syntax and Accent Placement: Towards a rule system for the derivation of pitch accent patterns in Dutch as spoken by humans and machines. PhD Dissertation, Leiden University.

Berinstein, A. E. (1979). A cross-linguistic study on the perception and production of stress. *UCLA Working Papers in Phonetics, 47*, 1–59.

Bolinger, D. (1958). A theory of pitch accents in English. *Word, 14*, 109–149.

Cohn, A. (1989). Stress in Indonesian and bracketing paradoxes. *Natural Language and Linguistic Theory, 7*, 167–216.

Cohn, A., & McCarthy, J. J. (1994). Alignment and parallelism in Indonesian Phonology. ROA-25.

Ebing, E. F. (1997). *Form and function of pitch movements in Indonesian.* CNWS publications 55. Leiden University.

Eefting, W. Z. F. (1991). The effect of information value and accentuation on the duration of Dutch words, syllables and segments. *Journal of the Acoustical Society of America, 89*, 412–414.

Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech, 1*, 126–152.

Goedemans, R. W. N., & van Zanten, E. (to appear). Stress and Accent in Indonesian. In David Gil (Ed.), *Malay/Indonesian Linguistics*. Asian Linguistics Series. London: Curzon Press.

Grijns, C. D. (1991). *Jakarta Malay; A multidimensional approach to spatial variation.* Leiden: Koninklijk Instituut voor Taal-, Land- en Volkenkunde [Royal Institute of Linguistics and Anthropology].

Halim, A. (1974). *Intonation in relation to syntax in Bahasa Indonesia*. Jakarta: Djambatan.

Hart, J. 't, Collier, R., & Cohen, A. (1990). *A perceptual study of intonation. An experimental-phonetic approach to speech perception.* Cambridge: Cambridge University Press.

Hayes, B. P. (1995). *Metrical stress theory: Principles and case studies*. Chicago: The University of Chicago Press.

Heuven, V. J. van (1994). Introducing prosodic phonetics. In C. Odé & V. J. van Heuven (Eds.), *Experimental studies of Indonesian prosody* (Semaian 9) (pp. 1–26). Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, Leiden University.

Heuven, V. J. van, & Broecke, M. P. R. van den (1982). Interactions of spectral distribution and decay time in the auditory perception of duration and their relevance to speech. *Progress Report of the Institute of Phonetics Utrecht, 7*(1), 21–41.

Heuven, V. J. van, & Zanten, E. van (1997). Effects of substrate language on the localization and perceptual evaluation of pitch movements in Indonesian. In C. Odé & W. Stokhof (Eds.), *Proceedings of the 7th International Conference on Austronesian Linguistics* (pp. 63–79). Amsterdam-Atlanta: Rodopi.

Hollander, J. J. de (1984). *Pedoman bahasa dan sastra Melayu*. Jakarta: Balai Pustaka.

Ikranagara, K. (1980). *Melayu Betawi grammar*. NUSA; Linguistic Studies of Indonesian and Other Languages in Indonesia, 9.

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

Laksman, M. (1994). Location of stress in Indonesian words and sentences. In C. Odé & V. J. van Heuven (Eds.), *Experimental studies of Indonesian prosody* (Semaian 9) (pp. 108–139). Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, Leiden University.

Lapoliwa, H. (1981). *A generative approach to the phonology of Bahasa Indonesia*. Pacific Linguistics, Series D-36. Canberra: The Australian National University.

Moeliono, A. M., & Dardjowidjojo, S. (1988). *Tata Bahasa Baku Bahasa Indonesia*. Jakarta: Balai Pustaka.

Muhadjir (1981). *Morphology of Jakarta dialect, affixation and reduplication*. NUSA, 11.

Nababan, P. W. J. (1981). *A Grammar of Toba Batak*. Pacific Linguistics D-37. Canberra: The Australian National University.

Nooteboom, S. G. (1972). Production and perception of vowel duration: A study of durational properties of vowels in Dutch. Unpublished PhD Dissertation, Utrecht University.

Odé, C. (1994). On the perception of prominence in Indonesian. In C. Odé & V. J. van Heuven (Eds.), *Experimental studies of Indonesian prosody* (Semaian 9) (pp. 27–107). Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, Leiden University.

Poedjosoedarmo, S. (1982). *Javanese influence on Indonesian*. Pacific Linguistics Series D-38. Canberra: The Australian National University.

Prentice, J. (1994). Manado Malay: Product and agent of language change. In T. Dutton & D. T. Tryon (Eds.), *Language Contact and Language Change in the Austronesian World.* Trends in Linguistics 7, 411–441. Berlin/New York: Mouton de Gruyter.

Ras, J. J. (1982). *Inleiding tot het Modern Javaans*. den Haag: M. Nijhoff.

Samsuri (1971). *Ciri-ciri prosodi dalam kalimat Bahasa Indonesia.* Flores: Nusa Indah.

Samsuri (1978). Fokus dan alat-alat pembentukannya dalam Bahasa Indonesia. Unpublished conference paper.

Steinhauer, H. (1980). On the history of Indonesian. In A. A. Barentsen et al. (Eds.), *Studies in Slavic and general linguistics*, Vol. 1 (pp. 349–375). Amsterdam: Rodopi.

Steinhauer, H. (1994). The Indonesian language situation and linguistics: Prospects and possibilities. Bijdragen Koninklijk Instituut voor Taal-, Land- en Volkenkunde [Royal Institute of Linguistics and Anthropology] 150-IV, 755–784.

Teeuw, A. (1984). *Leerboek Bahasa Indonesia* (4th ed.). Groningen: Wolters-Noordhoff.

Wallace, S. (1976). Linguistic and social dimensions of phonological variation in Jakarta Malay. Unpublished PhD Dissertation, Cornell University.

Woodrow, H. (1951). Time perception. In S. Stevens (Ed.), *Handbook of experimental phonology*. New York: Wiley & Sons.

Zanten, E. van (1994). The effect of sentence position and accent on the duration of Indonesian words: A pilot study. In C. Odé & V. J. van Heuven (Eds.), *Experimental studies of Indonesian prosody* (Semaian 9) (pp. 140–180). Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, Leiden University.

Zanten, E. van, & Heuven, V. J. van (1997). Effects of word length and substrate language on the temporal organisation of words in Indonesian. In C. Odé & W. Stokhof (Eds.), *Proceedings of the 7th International Conference on Austronesian Linguistics* (pp. 201–213). Amsterdam-Atlanta: Rodopi.

Zanten, E. van, & Heuven, V. J. van (1998). Word stress in Indonesian: Its communicative relevance. Bijdragen Koninklijk Instituut voor Taal-, Land- en Volkenkunde [Royal Institute of Linguistics and Anthropology] 154-I, 129–147.

Zanten, E. van, & Heuven, V. J. van (to appear). Word stress in Indonesian: Fixed or free? To appear in NUSA; Linguistic Studies of Indonesian and Other Languages in Indonesia.

Zubkova, L. G. (1966). Vokalizm Indonezijskogo jazyka. PhD Dissertation, Leningrad University.

Prosodic structure

Perceived prominence and the metrical-prosodic structure of Dutch sentences

Karijn Helsloot and Barbertje M. Streefkerk
University of Amsterdam

1. Introduction

In the past ten to fifteen years, different prosodic parsing models have been advanced for the purpose of text-to-speech (TTS) systems. But a near-natural speech sound realisation of written texts continues to be an extremely difficult matter. With this paper, we do not intend to resolve the existing problems, but to provide new information on the subject which may lead to improvements of existing TTS systems.

In comparison with previously proposed prosodic parsing models for Dutch TTS systems (cf. Baart 1987; Dirksen & Quené 1993; and Quené & Kager 1993), we claim that our proposal is easier to implement, and at the same time more in accordance with the metrical variation that is observed in natural sentences. Instead of a set of metrical principles translating syntactic phrase structure rules, our model indirectly reflects syntactic relations by assigning different metrical values to verbs, nouns and modifiers. This metrical variation, in turn, combines with a set of prosodic well-formedness constraints, giving rise to four degrees of prominence, each of which has to be translated into appropriate acoustic values. Previous models recognise only two degrees of prominence, viz. zero prominence, and accent prominence.

Since the corpus on which our analysis is based consists of newspaper sentences presented out of their context, the semantic-pragmatic distinction between given and new information is not included in our metrical parsing constraints.

It is generally assumed that prosodic parsing includes accentuation as well as phrasing. The perception task on the basis of which we define the metrical constraints, is restricted to information about accentuation (prominence), however. No constraints are thus formulated that refer to (different strengths of) boundaries.

2. The phonetic material

The speech material is selected from the Dutch Polyphone corpus. This corpus consists of 12500 newspaper sentences. A total of 5000 speakers were asked to read five sentences, to be recorded over the telephone (for more details see Damhuis et al. 1994). From this corpus, we took a random set of 50 sentences for the purpose of a metrical-prosodic analysis. Although the grammatical structure of the sentences varies, they are all declarative.

2.1 The corpus of Dutch sentences

The 50 newspaper sentences consist, on average, of 10.4 words per sentence, and the average number of syllables per sentence is 18.5, as shown in Table 1. About half of the words are function words and the other half are content words. As expected, function words are perceived as being less prominent than content words.

2.2 Listening experiment

The 50 sentences are part of a much larger set of sentences selected by Streefkerk for the purpose of a study on prominence perception (Streefkerk 1997, 2002). A first perception experiment, executed by Streefkerk, involved 500 sentences produced by 50 male and 50 female speakers. Ten listeners, all students

Table 1. Number and means of words and syllables over 50 sentences and per sentence.

	Total number	Mean per sentence
words	519	10.4
content words	**278**	**5.6**
function words	**233**	**4.7**
rest words	**8**	–
syllables	924	18.5

Table 2. Example of the results of the listening experiment. The table shows the cumulative judgements of the listeners and the resulting degrees of prominence.

Listener #	De *the*	vliegtuigkaping *hijacking*	werd *was*	tijdens *during*	de *the*	vlucht *flight*	opgelost *resolved*
1	0	1	0	1	0	1	0
.	.	.	.	.	.	.	.
9	0	1	0	1	0	0	1
10	0	1	0	1	0	1	0
Sum first	0	8	0	8	0	4	1
Sum second	0	8	0	8	0	6	3
Sum total	**0**	**16**	**0**	**16**	**0**	**10**	**4**

from the Humanities Faculty at the University of Amsterdam, were asked to indicate which words were realised with emphasis. The 500 sentences were presented in 4 random order sessions, which differed per listener, to compensate for possible learning effects. The first two sessions contained 150, and the last two sessions contained 125 sentences. The perception experiment was performed on a UNIX workstation, and the results of each listener were automatically stored. While hearing the sentence through closed headphones, the listeners saw the sentence on a monitor. Under each word, on the monitor, a button was placed. The subjects had to click on the button when a given word was perceived as being spoken with emphasis. To test the consistency of the listeners, 50 sentences were presented twice to each listener. This set of 50 sentences is used for the metrical-prosodic analysis.

An example of the perception results for one sentence is given in Table 2. The sentence *De vliegtuigkaping werd tijdens de vlucht opgelost.* 'The airplane hijacking was resolved during the flight' was scored twice by the 10 listeners. The 20 judgements for each word were totalled, yielding a score between 0 (no mark) and 20 (all listeners marked this word twice as emphasised). We assume that the resulting scale of judgements is an indication of the degree of prominence involved: the higher the score the more prominent a given word is.

It deserves mention that the listeners differ quite remarkably with respect to the number of emphasised words they perceive in one and the same sentence. While some listeners assign a mean number of 1 prominence per sentence (see e.g. listener 7, Table 3), others assign 4 prominences (see e.g. listener 9, Table 3). It is striking, for instance, that there were only four instances in the corpus where all listeners agreed that a certain word is emphasised, yielding 20 marks. These facts argue in favour of a relative, rather than an absolute, metrical representation. That is, a prosodic analysis which rigidly translates the syntactic sur-

Table 3. Number of prominence judgements per listener after first and second parsing.

Listener #	1	2	3	4	5	6	7	8	9	10	sum
First 50	71	50	160	165	135	132	50	109	156	172	1200
Second 50	71	51	165	202	130	211	50	149	209	158	1396

Table 4. Distribution of prominence judgements between 0 and 20.

Prominence	Total	Mean per sentence
11–20	136	2.7
1–10	139	2.8
0	244	4.9
Total	519	10.4

face structure into prosodic constituents, as proposed for instance by Nespor and Vogel (1986), leads to an abundance of prosodic heads and boundaries which have no acoustic and perceptional counterparts.

Table 3 also provides evidence of a learning effect. Listeners 4, 6, 8, and 9 marked substantially more words as prominent during the second parsing than during the first one. That is, in the second parsing words are added to the set of prominent words marked during the first parsing (see Streefkerk & Pols 1998). Although the differences within and between listeners are significant, we still consider their cumulative judgements to be a useful alternative for prominence labelling of the speech material.

In Table 4 we present the distribution of prominence marks according to a threefold distinction: (i) 0 marks, (ii) from 1 to 10 marks, and (iii) from 11 to 20 marks. The mean number of 0 marked words is 4.9 per sentence. This equals the amount of function words per sentence, as illustrated in Table 1. The sum of marked words gives rise to a mean of 5.5 words per sentence. This equals the amount of content words per sentence (see Table 1).

A final observation must be made here with regard to the listening experiment. Since the listeners were asked to indicate emphasised words, and not emphasised syllables, and they were *not* asked to indicate degrees of emphasis, the results do not give us information about (i) the location of the emphasis within the word, and (ii) the presence of weakly stressed words/syllables. With respect to (i), in general, the lexically stressed syllable of the word is also the syllable actually realised with prominence. Only in a very few cases, can lexical stress shifts be observed. With respect to (ii), unfortunately the prominence values of polysyllabic function words and of secondary stresses within relatively long content words were not perceptually tested. But in a similar experiment on a different set of sentences from the Polyphone corpus, the task included perceived prominence of words versus perceived prominence of syllables. The listeners indicated a mean number of 2.9 words, but a mean number of 5.1 syllables per sentence as being prominent (Streefkerk et al. 1997). This result suggests that weakly stressed syllables are indeed perceived when the perception task is formulated differently.

Other perception tests in which degrees of prominence were asked for, similarly indicate that listeners are able to differentiate between unstressed, weakly stressed and strongly stressed syllables (see Helsloot 1993, 1995). In addition, the acoustic signals as well as careful listening to short stretches of the Dutch sentences reveal the presence of different stress degrees.

3. The metrical-prosodic analysis

With the results of the listening experiment on the one hand, and rather elaborated theories of prosodic phonology on the other (cf. among others, Selkirk 1980, 1986), we hypothesised that a mapping of the two would be feasible. The six-level organisation assumed by prosodic phonology, i.e. the syllable, the foot, the prosodic word, the phonological phrase, the intonation phrase and the phonological utterance, appeared to be a far too rich as well as a far too rigid system, however. Instead of six levels, it is possible to distinguish only four levels. Moreover, the assignment of prosodic constituent structure to sentences based on morphological and syntactic information gives rise to heads and constituents which, very often, are not encountered in the perception results. A relational-based metrical grid representation, as initially proposed by Liberman and Prince (1977), extended by restrictions on the number of hierarchical levels, instead, allows for a more adequate analysis of the metrical structure of the material.

3.1 Prosodic input constraints

Since we are dealing with sentences read aloud, not with spontaneous speech, almost all syllable inputs are properly realised in the output. Syllable deletion or syllable insertion, well-known phenomena in Dutch spontaneous speech (see Kuijpers & van Donselaar 1998), occur in just a few cases. In the 50 sentences, we found four instances of syllable deletion, and two instances of syllable insertion. The vowel involved is always Schwa:

(1)

syllable deletion		*syllable insertion*	
gist[∂]ren 'yesterday'	> gistren	merkt	> mer[∂]kt '(he) notices'
vriend[∂]lijke 'friendly'	> vrienlijke	half uur	> hall[∂]f 'half an hour'
verzek[∂]ringsagent 'insurance agent'	> verzekringsagent		
Ned[∂]rland 'the Netherlands'	> neland		

Generally, in a sequence of two or more unstressed syllables, the left-most Schwa tends to be deleted, and in pre-boundary position, or in a sequence of adjacent stresses, a Schwa is inserted if permitted by the segmental environment. In other words, rhythmic lapses and clashes are possibly resolved at the syllable level. Obviously, a TTS system must include these rhythmically-driven syllable deletions and insertions.

Apart from these insertions/deletions, input syllables are realised at the surface. This observation is metrically represented by a mark on the lowest metrical grid level:

(2) *Syllable Constraint*
All syllables receive a level-1 mark on the metrical grid.

Thus, the sentence in (3) is initially parsed into a sequence of level-1 marks:

```
(3)   x x       x    x    x      x    x    x    x  x
     D e  o n t s p o o r d e  t r e i n  r a m d e  h e t  t a l u u d.
     'The derailed train rammed the bank'
```

3.2 Function words

Our corpus contains a total of 233 function words, comprising Determiners, Auxiliaries/Modals/Copulas, Prepositions, Possessive Pronouns, Complementizers, Personal Pronouns, Reflexive Pronouns, (Anaphoric) Demonstrative Pronouns, and Conjunctions. Of these function words, 216 are monosyllabic, 17 are polysyllabic. Table 5 presents the relevant distributions.

The listeners perceived 205 monosyllabic function words as bearing no prominence at all. Eleven monosyllables were perceived as bearing a (very) low degree of prominence ($1 < \text{Prom} \leq 10$). Three distinct explanations can be given for the perceived prominence, albeit low, of these function words: (i) the monosyllabic function word occurs in absolute sentence-initial position;

Table 5. Number and means of monosyllabic and polysyllabic function words regarding prominence degrees.

Function words	Total	0 Prom	%	$1 < \text{Prom} \leq 10$	%	$10 < \text{Prom} \leq 20$	%
Monosyllabic	216	205	95	11	5	–	–
Polysyllabic	17	12	70	2	12	3	18

(ii) the monosyllable is prominent in order to avoid a rhythmic lapse; and (iii) the monosyllable receives prominence because it is part of a slowly read speech string in which all syllables are realised with prominence. Some examples (the numbers following the sentence fragments correspond to the number of marks assigned by the listeners to the separate words):

(4)	i.	**Ik** was hier al	**1**.6.2.3
		'I was here already'	
		In feite is Nederland	**1**.19.0.2
		'In fact the Netherlands is'	
	ii.	… bevatten **dat** ze had gewonnen	0.**1**.0.0.4
		'… comprehend that she had won'	
		… wachten **op** het moment …	1.**1**.0.10
		'… wait for the moment'	
	iii.	van algemene rouw **ten** gevolg	0.11.14.**1**.2
		'as a result of general mourning'	

The explanations (i) and (iii) are non-grammatical in nature, and possibly also speaker-dependent. It is certainly not the case that sentence-initial monosyllabic function words tend to be realised with prominence, or are systematically perceived as being stressed. A decreased speech rate, observed in a few readings, is also not grammatically determined. That is, slow reading must not be incorporated in the basic prosodic parsing of a TTS model. In contrast, explanation (ii), like the above-mentioned phenomena of syllable deletion/insertion, provides an example of rhythmic readjustment: a stress is added in order to avoid a rhythmic lapse. In Section (4.2.2), this grammatically determined rhythmic readjustment is formalised in terms of a prosodic output constraint.

As is indicated in the table, 95% of the monosyllabic function words are perceived as completely unstressed. This amounts to the following TTS prosodic input constraint:

(5) *Function Word Constraint I:*
Monosyllabic function words do not receive a grid mark on level-2 or higher.

Regarding polysyllabic function words, the following constraint is proposed.

(6) *Function Word Constraint II:*
The head syllable of a polysyllabic function word receives a level-2 mark in the metrical grid.

Although the table indicates that 70% of all polysyllabic function words in the corpus were perceived as bearing no prominence at all, we maintain that polysyllabic function words are more prominent than monosyllabic ones. First of all, in citation form, a speaker of Dutch will identify one of the syllables of a functional polysyllable as bearing word stress. Secondly, as stated in the introduction, due to the formulation of the perception task, listeners focussed on words realised with emphasis. Clearly, function words are generally not realised with emphasis, although they may be realised with a low prominence degree. And thirdly, the acoustic representations of the sentences clearly indicate that weakly stressed syllables have particular acoustic properties. Although an acoustic analysis is outside the scope of this paper, a TTS model should translate weakly stressed syllables in order to get close-to-natural realisations. In the models for Dutch known to us, weak or secondary stress is completely neglected.

The polysyllabic function words perceived as being realised with prominence are mostly emphasised prepositions. In other words, the prepositions received contrastive focus:

(7) a. De vliegtuigkaping werd *tijdens* de vlucht opgelost.
'The hijacking was resolved *during* the flight' (and not *after* the flight)
b. Mijn verzekeringsagent woont *tussen* de medisch specialisten in Beugen.
'My insurance agent lives *amid* the medical specialists in Beugen' (and not in another neighbourhood, as you would expect)

Obviously, this contrastive focus must be accounted for by another constraint than *Function Word Constraint II*.

3.3 Content words

Content words are stressed, function words are not. This prosodic distinction is uncontroversial in the phonological literature, as well as in the TTS models that have been proposed for Dutch, and for other stress-based languages (see e.g. O'Shaughnessy 1976; Baart 1987). Indeed, in our corpus of read sentences, listeners perceived prominence on many if not all nouns, verbs, adjectives and

Table 6. Number and means of content word categories regarding prominence degrees.

Content words	Total	0 Prom	%	1 < Prom ≤ 10	%	10 < Prom ≤ 20	%
Nouns	143	2	1,3	34	23,7	107	**75**
Verbs	50	19	**38**	14	29	17	33
Adverbs	33	–		11	33	22	**67**
Adjectives	46	–		7	15	39	**85**

adverbs. In Table 6, the exact results are given. Adverbs and adjectives are always perceived as being realised with prominence. With two exceptions, nouns, too, are perceived as prominent (1,3% *vs*. 98,7%). This is also true of most verbs, although the rates for verbs are more balanced (38% no prominence vs. 62% prominence).

This prosodic property of content words leads to the following constraint:

(8) *Content Word Constraint I:*
The head syllable of a content word (either monosyllabic or polysyllabic) receives a level-2 mark on the metrical grid.

As shown by the figures in Table 6, verbs are prosodically less prominent than the other word categories.[1] This tendency is characteristic of west-Germanic languages like English and Dutch. As reported by Baart (1987:57), scales of accentability for content word classes in English place main verbs lower than nouns, adjectives and adverbs (cf. Lea 1979; O'Shaughnessy & Allen 1983). For Dutch, Kruyt (1985) argues that verbs have a lower accentability degree than nouns and adjectives, but a slightly higher degree than adverbs. In our corpus, the position on the scale of accentability for Dutch adverbs, as proposed by Kruyt, cannot be confirmed. Adverbs are mostly perceived as highly prominent. The following constraint formalises the findings, as reported in Table 6:

(9) *Content Word Constraint II:*
The head syllable of a noun, adjective or adverb receives a level-3 mark on the metrical grid.

The constraints proposed so far assign the following metrical grid representation to the sentence in (10).

(10)
```
           x          x                          x
           x          x      x                   x
 x   x     x     x    x      x    x    x    x    x
 De  ont   spoor de   trein  ram  de   het  ta   luud.
```
'The derailed train rammed the bank.'

Before addressing deviations from *Content Word Constraints I* and *II*, we first address the metrical input representation for compounds and complex verbs.

3.3.1 *Compounds*

Dutch has a highly productive system of compounding. Almost a fifth of all content words in our corpus consist of compounds. Although the number of nominal compounds exceeds by far the number of adjectival, verbal and adverbial ones, all four categories allow for productive compounding. The lexical or morpho-phonological properties of Dutch compounds have given rise to a variety of metrically rich and complex representations (see Visch 1989; Booij 1995). We suggest that the metrical representations of compounds at the phrase-level, however, can be reduced to two. The first type is lexically characterised by two adjacent stressed syllables, the second type by two non-adjacent stressed syllables. All disyllabic compounds belong to the former type. Metrically larger compounds belong to either the former or the latter type:

(11) *Type I Compounds* *Type II Compounds*

```
x             x             X                   x
x  X       >  x             X     x         >   x       x
x  X  (x)     x   x   (x)   X  x  x   (x)       x   x   x   (x)
```

wíelrenner	(N)	ménsenschùw	(N)
‘cyclist’		‘shy’	
kléurrijk	(A)	èensgezínd	(A)
‘colourful’		‘unanimous’	
ópzien	(V)	úitgerùst	(V)
‘look up’		‘rested’	

Type I compounds thus behave like simplex content words. A TTS model just needs to know which of the two syllables is the head of the word. Whether the word forms a compound or not is irrelevant for the purpose of an adequate metrical TTS representation. In (12) a sentence is provided containing both a type I and a type II compound.

(12)

```
    x               x           x                     x
    x               x  x        x       x             x
x   x    x  x    x  x  xx  x  x x  xx   x    x    x   x  x
De  werknemers   van  Unilever  in India hebben het moeilijk.
```

De <u>werknemers</u> van <u>Unilever</u> in India hebben het moeilijk.
‘The employees of Unilever in India are in trouble.’

3.3.2 *Complex verbs*

Dutch is rich in complex verbs, formed by a verbal stem preceded by a prepositional particle. When inflected, the verb and particle are separated from each other, such that the particle occurs in clause-final position and the verbal stem in the second position of the main clause. In participle constructions and infinitives, particle and stem form a compound, except in infinitive constructions in which the infinitive marker *te* occurs. In the latter case, the word order is particle+*te*+verbal stem. In other words, there are two constructions leading to *surface compounds* and two constructions leading to *surface separation*.

Lexically, most verbal compounds have main stress on the particle: lángskomen 'come by', óphalen 'fetch'. With respect to prominence perception, the following tendencies are observed in our corpus: (i) the *surface compounds* are always perceived as prominent (e.g. (13a)); (ii) the verbal stems in the separated forms are never perceived as being prominent (e.g. (13b)); (iii) the particles in the separated forms are always perceived as prominent in the particle+*te*+verb infinitive constructions (e.g. (13c)); but (iv) the clause-final particles in inflected forms are only perceived as prominent if they do *not* occur in sentence-final position, i.e. when they occur in the non-final clause of the sentence (e.g. (13d)); in sentence-final position they are not perceived as prominent (e.g. (13b)). (Verbal stem and particle are marked by italics, and the prominent word by boldface).

(13) a. Wij hebben de beukeboom die in de tuin stond, ***omgehakt***.
'We have cut down the beech-tree, which stood in our garden.'
b. Ajax *ging* in de kwartfinale van het eindtoernooi kansloos *onderuit*.
'Ajax did not stand a chance in the quarter-finals of the final tournament.'
c. De voetballer is verplicht om zijn contractuele verplichtingen ***na*** *te komen*.
'The football player is obliged to observe his commitments by contract.'
d. Na morgen *zet* het vriendelijke weer zich ***voort*** en zijn de buien verdwenen.
'After tomorrow the friendly weather proceeds and the showers disappears.'

On the basis of these heterogeneous findings it is not clear how to represent complex verbs metrically.

3.4 Lexical modifiers

A final metrical input constraint will be presented now. The results of the perception task show that in most cases in which an argument is modified by a word (and not by a phrase), this modifier is perceived as bearing more prominence than the modified word. Relevant sequences are adjective-noun, verb-adverb, adverb-verb, and adverb-adjective sequences. A couple of examples are given below.

(14)	kwalit**ei**tsvolle m**a**rathonloper	11.6	Adj-N
	'(a) marathon runner of high quality'		
	vanuit r**ij**dende **au**to's	0.12.1	Adj-N
	'from moving cars'		
	risk**ee**rde bew**u**st	0.19	V-Adv
	'risked consciously'		
	v**aa**k m**ee**t men	10.0.0	Adv-V-Pro
	'it is often measured'		

Of 35 adjective-noun sequences in the corpus, 30 were perceived as having a higher degree of prominence on the adjective than on the noun. This amounts to 86% of all adjective-noun sequences. In the case of adverb-verb or verb-adverb sequences, there are no instances at all in which the verb is perceived as bearing a higher degree of prominence than the adverb. On the basis of these facts, we formulate the *Modifier Constraint*, which assigns a level-4 mark to a lexical modifier.

(15) *Modifier Constraint*: Each lexical modifier receives a level-4 mark.

Regarding the 5 instances in which the adjective has no higher degree of prominence than the noun, it is observed that in 4 instances the prominence degrees of the two words are either identical or almost identical:

(16)	a.	potenti**ë**le betr**o**kkenen	9.9
		'potentially concerned (persons)'	
	b.	versch**i**llende **au**to's	8.9
		'different cars'	
	c.	om h**a**lf z**e**s	0.6.8
		'at half past five'	
	d.	**a**lgemene r**ou**w	11.14
		'general mourning'	

The only clear-cut iambic, or weak-strong pattern is realized on:

(17) van de **ou**de **fie**tsroutes 0.0.3.12
‘of the old cycling routes’

Focus, lexicalisation, frequency, non-finality, and maybe other explanations might be advanced as underlying this marked pattern, but all these properties are equally characteristic of one or more of the adjective-noun sequences with a trochaic pattern. The tendency to emphasise the adjective and not the noun clearly prevails. In fact, the phrases in (16) and (17) also allow for a trochaic, or strong-weak realisation.

Negative particles and deictically used demonstrative pronouns are similarly realised with a high degree of prominence. These elements too are subject to the *Modifier Constraint*.

The entire set of metrical input constraints, applied to the sentence provided earlier in (10), gives rise to the following representation:

(18)

```
                x
                x        x                      x
                x        x     x                x
x    x          x    x   x     x    x    x   x  x
De  ontspoorde trein ramde het taluud.
```

‘The derailed train rammed the bank.’

4. Prosodic output constraints

The metrical input constraints do not account for all the metrical patterns observed in our corpus. Two metrical patterns in particular require a specific account, necessitating reference to the beginnings of sentences and the endings of sentences. Sentence-initial content words (except verbs) have a higher degree of prominence than following content words. In the case of sentence-initial adjective-noun or adverb-verb sequences no intervention is needed, since lexical modifiers receive a higher grid mark by means of the *Modifier Constraint* presented in Section 3.4. But when the first constituent contains a non-modified noun and the immediately following constituent is either a modifier or a non-modified noun, ill-formed metrical patterns arise. That is, the sentence-initial noun has either a lower grid-level or the same grid-level as the following content word:

(19)

```
      x
 x    x              x    x
 x    x              x    x
 x    x              x    x
([N] [AN]          ([N] [N]
```

The examples in (20) illustrate the actually observed prominence patterns. The prominence degrees refer to the words in italics:

(20)	a.	De *bewijslast* tegen de *verdachte*. 'The burden of proof against the suspect'.	19.16
	b.	De *verzekeraars* hebben het *moeilijk*. 'The insurance companies are in trouble'.	17.17
	c.	Zijn *reputatie* als *kwaliteitsvolle marathonloper*. 'His reputation as a marathon runner of high quality'.	16.11.6

To ensure that a higher degree of prominence is realised in sentence-initial position than on the following content word, the *Sentence-Initial Constraint* is assumed to operate:

(21) *Sentence-Initial Constraint:*
The first level-3 mark in the sentence receives a level-4 mark.

The explicit reference to words with a level-3 grid mark prevents verbs to be subject to this constraint. In fact, a sentence like *Hij eet een appel* 'He eats an apple' has the highest prominence degree on the noun *appel* and not on the verb *eet*.

Combinations of first name and surname, title and surname, and dates, constitute exceptions to this pattern, insofar that the highest prominence is on the second content word:

(22)	P**a**ter Gr**oe**newegen 'Father Groenewegen'	4.15
	in de maand februari 'in the month of February'	0.0.0.17

Sentence-finally, a similar trochaic pattern is observed. That is, if the input gives rise to a strong-strong sequence (level-3 plus level-3, or level-3 plus level-4), a strong-weak output is realised. Since sentence-final words are typically perceived as weakly prominent, the relevant constraint is as follows:

(23) *Sentence-Final Constraint*:
The final level-3(4) mark in the sentence is deleted.

The fact that the base position for verbs in Dutch is sentence-final, causes many sentences in our corpus to be produced with a sentence-final trochaic pattern. Verbs have a low grid level, as illustrated by the sentence-final strings in (24) (the prominence marks refer to the words in italics):

(24) a. aan het *orgel gewijd* 14.2
'dedicated to the organ'
b. een *shuttle gelanceerd* 15.0
'a shuttle launched'
c. en zijn de *buien verdwenen* 10.1
'and have the showers gone'

The constraint is required, however, in order to account for outputs like those in (25):

(25) a. bij een *winkelcentrum* te *Houten* 11.3
'at a shopping centre in Houten'
b. op het *moment* van *gloreren* 10.5
'at the moment of glory'
c. vanuit *rijdende auto's* 12.1
'from moving cars'

The Sentence-Final Constraint reduces the final strong prominence to a weak prominence by grid-mark deletion, as indicated in (26) by the angled brackets:

```
(26)  x   <x>
      x    x
      x    x
```

An example is given below.

```
(27)                x           <x>
                    x            x
      x   x     x   x    x     x x x
      op  het  moment   van  gloreren
```

'at the moment of glory'

4.1 Prosodic maximality

Violations of the *Sentence-Final Constraint*, yielding an iambic pattern instead of the expected trochaic pattern, do occur, however. Of the 50 metrically analysed sentences, 7 have an iambic ending. Rhythmic alternation and domain maximality are the underlying reasons here. To start with the latter phenomenon, domain maximality, a sentence-final iamb is created when the preceding content word is 'too far away' from the final word. More precisely, the trochee must be realised within a maximum number of syllables. If the number of syllables intervening between the head syllables of the final and pre-final content word exceeds four, the input pattern is left unchanged. Two examples are given below (head syllables are given in boldface; the prominence marks refer to the words in italics):

(28) a. nauwelijks *be**vat**ten dat ze had ge**won**nen* 0.4
‘hardly comprehend that she had won’
b. bij de *be**vol**king tegen de poli**tiek*** 9.14
‘among the population against politics’

Five syllables occur between the two head syllables *vat* and *won* in (28a), six between the head syllables *vol* and *tiek* in (28b). Helsloot (1995) dedicates a chapter to prosodic phrase maximality in Italian. Roughly, in Italian, the maximal phrase includes six to seven syllables, parsed into three feet. The final parts of the strings in (28) give rise to such maximal phrases: [dat ze had gewonnen] and [tegen de politiek]. In fact, in addition to the sentence-final verb *gewonnen* the complementizer *dat* is perceived by the listeners as having low prominence, and in (b), the disyllabic preposition *tegen*, although not perceived as bearing prominence, is expected to be realised with a low degree of prominence, on the basis of our *Function Word Constraint II*. The trochaic pattern required by the *Sentence-Final Constraint* cannot be realised because of maximality conditions on metrical phrasing. In the prosodic framework we are presenting here, phrases and boundaries are not explicit entities. However, the observed violation of the *Sentence-Final Constraint* can also be accounted for by a constraint on rhythmic alternation, which refers to alternating stress degrees instead of phrase size.

4.2 Rhythmic alternation

Prosodic well-formedness mainly involves rhythmic alternation: at the phrasal level stresses must be distributed in accordance to principles of recurrence and hierarchical ordering (cf. amongst others, Prince 1983; Hayes 1984; Halle & Vergnaud 1987; Helsloot 1995). Recurrence is generally assumed to be of a binary nature: a strong beat is followed by a weak beat. This alternation applies at hierarchically ordered levels of organisation. In other words, binary alternation is observed from the smallest to the largest domains of prosodic organisation. Metrical grids insightfully represent this organisation. Well-formed rhythmic alternation is shown in (29a), ill-formed alternations in (29b) and (29c).

```
(29) a. x                   b.    x                  c.
        x       x                 x   x                  x         x
        x   x   x                 x   x                  x         x
        x x x x x x             x x x x x x              x x x x x x
```

4.2.1 *Clash avoidance*

Obviously, the words of a language when grouped into phrases do not always give rise to well-formed rhythmically alternating patterns. The input may show up with a so-called stress clash (as is the case in (29b)). Since stress clashes are hard to produce physically, they generally do not occur in the actual realisation of the phrase. A clash can be resolved by means of (i) deletion of one of the stresses, (ii) movement of one of the stresses, or (iii) lengthening of the interval between the stresses. Whether one resolution or the other is selected depends on the linguistic context. In Dutch, all three resolutions are found. It is the rightmost stress (i.e. the second stress) which is most frequently affected. Where a clash exists between three stresses, things are even more complex. In our corpus, about ten instances are found in which a sequence of three content words gives rise to a double stress clash.

In Section 4.2, we referred to instances violating the *Sentence-Final Constraint*. Instead of a final trochee, a iamb is realised. The violation is due to stress-clash avoidance. Consider the metrical input representation of the sentence in (30).

(30)

```
                               x
    x                          x          x        x
    x         x                x          x        x       x
x   x   x     x     x   x  x   x  x    x  x   x    x     x x
De rechter spreekt van een onthutsend gebrek aan plichtsbesef.
```

'The judge speaks of a bewildering lack of sense of duty.'

Application of the *Sentence-Final Constraint* would delete the sentence-final level-3 mark, i.e. the third mark on *plicht*. However, the perceived prominence pattern in the listening experiment is 5 on *gebrek* and 11 on *plichtsbesef*, a clear iambic pattern. The representation in (30), in fact, shows up with two stress clashes, as indicated by the boldfaced marks.

On the basis of the perceived prominences, we propose a *Clash Resolution Constraint* which requires left-headedness at the highest level of organisation.

(31) *Clash Resolution Constraint:*
Create left-headed alternating grid representations.

```
x   x   x       x
x   x   x   >   x       x
x   x   x       x   x   x
```

On the basis of this constraint, the above sentence is represented as follows:

(32)

```
   x                           x
   x                           x        <x>        x
   x          x                x         x         x       x
x  x   x      x     x   x  x   x  x    x x   x     x     x x
De rechter spreekt van een onthutsend gebrek aan plichtsbesef.
```

'The judge speaks of a bewildering lack of sense of duty.'

That is, the level-3 mark on *gebrek* will not be realized in the output.

The two relevant constraints are ranked with respect to one another: the *Clash Resolution Constraint* dominates the *Sentence-Final Constraint*, which explains the fact that the latter is violated.

Other examples from our corpus subject to the *Clash Resolution Constraint*, but not in sentence-final position, are given in (33).

(33) a.

```
        x   <x>
        x   <x>     x
 x      x    x      x
 x  x   x    x  x   x   x
zitten vol groene vruchten          0.5.3.8
```

'are full of green fruits'

b.

```
    x       <x>    <x>
    x       <x>     x
    x        x      x       x
 x  x   x    x  x   x  x    x
aanzienlijk minder mensenschuw      16.1.4
```

'considerably less shy'

The *Modifier Constraint* assigns a level-4 mark to the modifiers *vol* and *groene* in (33a), and *aanzienlijk*, *minder* and *mensenschuw* in (33b). Since these level-4 marks are not separated from each other by intervening level-3 and level-2 marks, they are subject to the *Clash Resolution Constraint*. This latter constraint outranks the *Modifier Constraint*, i.e., it is satisfied at the expense of the *Modifier Constraint*.

4.2.2 *Lapse avoidance*

The rhythmic counterpart of the clash is the likewise ill-formed *lapse*, as in (29c) above. The metrical input may give rise to a relatively long sequence of syllables that are only characterised by level-1 marks. Such a pattern gives rise to a monotonous realisation. Generally, such patterns are avoided in natural speech. The *Lapse Resolution Constraint* accounts for this avoidance.

(34) *Lapse Resolution Constraint:*
In a sequence of more than three level-1 marks a level-2 mark is assigned to the central syllable with a full vowel.

In Section 3.2, we mentioned that monosyllabic function words are sometimes perceived by the listeners as being realised with a low degree of prominence. The avoidance of a rhythmic lapse explains this prominence (the mark between square brackets is added):

(35)

```
                                                                       x
     x         [x]                x              x       [x]           x
  x  x   x      x    x    x    x  x    x         x     x  x    x   x   x
bevatten       dat  ze  had  gewonnen          wachten  op  het  moment
```

'comprehend that she had won' 'wait for the moment'

Recall that listeners were asked to indicate which words in the sentences were realised with emphasis. Word-internal weakly stressed syllables were, therefore, not marked separately. Another kind of perception test is needed to obtain judgements about the presence or absence of weakly stressed syllables. However, TTS models which do not translate weak (or level-2) stresses in acoustic terms, will produce monotonous realisations of such weakly stressed intervals. The *Lapse Resolution Constraint* prevents the occurrence of such realisations.

5. Conclusions

A perception-based metrical-prosodic analysis of Dutch sentences which were read aloud has led to a metrical grid representation containing four distinct levels of prominence. The lowest level corresponds to the syllable level, the second to weakly stressed syllables – either a function word, a non-head syllable of a long content word, or the head syllable of a verb, the third to the (lexically stressed) head syllable of nouns, and the fourth to the head syllable of adjectives and adverbs, as well as to negative particles and deictically used demonstrative pronouns. This representation is accounted for by a set of prosodic input constraints: the Syllable Constraint, the Function Word Constraints I and II, the Content Word Constraints I and II, and the Modifier Constraint.

In addition, a number of sentence-level output constraints are formulated which account for the rhythmic well-formedness of the sentences: the Sentence-Initial Constraint, the Sentence-Final Constraint, the Clash Resolution Constraint, the Lapse Resolution Constraint.

The prosodic output constraints are often in conflict with the prosodic input constraints. The surface realisations indicate that the former are ranked higher in the constraint hierarchy than the latter. Output constraints themselves may also be in conflict with one another. For instance, the Sentence-

Final Constraint can be violated by the more highly ranked Clash Resolution Constraint.

The metrical grid representations resulting from the constraints match extremely well with the presence versus absence of prominences as perceived by the listeners: of the 275 words perceived as prominent, only eleven do not receive a proper metrical representation (see the prominence marks assigned to monosyllabic function words in Table 5), and only one word which receives a level-3 mark (or higher) in the metrical grid representation was not perceived as prominent at all. With respect to level-2, level-3 and level-4 marks on the one hand, and perception marks on the other, it is observed that the correspondences are relatively and locally manifested, but not absolutely. The fact that each sentence was read by just one speaker did not allow us to correct for speaker-dependent pronunciation. The next step is to verify the proposal on the basis of a different set of sentences, whereby first, the text-to-speech system automatically assigns the appropriate metrical grid representation to the sentences; second, the grid representation is translated in acoustic features; and finally, the synthesised sentences are evaluated by listeners.

Notes

1. The imperative form is the only verb form which in most languages is generally realized with a high degree of prominence (cf. Helsloot 1995). However, imperatives do not occur in our corpus.

References

Baart, J. (1987). Focus, Syntax and Accent Placement. Dissertation, Leiden University.

Booij, G. (1995). *The Phonology of Dutch*. Oxford University Press.

Damhuis M., Boogaart T., in 't Veld C., Versteijlen M., Schelvis W., Bos L., Boves L. (1994). Creation and analysis of the Dutch Polyphone corpus. *ICSLP 94* (p. 1803). Yokohama.

Dirksen, A., & Quené, H. (1993). Prosodic analysis: The next generation. In V. J. J. P. van Heuven & L. C. W. Pols (Eds.), *Analysis and Synthesis of speech: Strategic research towards high-quality TTS generation* (pp. 131–144). Mouton de Gruyter.

Dirksen, A., & Menert, L. (1997). *Fluent Dutch Text-to-Speech*, Version 1.0. Fluency Speech Technology, Utrecht.

Halle, M., & Vergnaud, J-R. (1987). *An essay on stress*. Cambridge, MA: MIT Press.

Hayes, B. (1984). The Phonology of Rhythm in English. *Linguistic Inquiry*, *15*, 33–74.

Helsloot, C. J. (1993). The Prosodic Word: an Endless Domain? *Proceedings, 17*, 129–146. Institute of Phonetic Sciences, University of Amsterdam.

Helsloot, C. J. (1995). Metrical Prosody. A Template-and-Constraint Approach to Phonological Phrasing in Italian. HIL Dissertation 16, HAG, Den Haag.

Kruyt, J. (1985). Accents from Speakers to Listeners. Dissertation, Leiden University.

Kuijpers, C., & van Donselaar, M. (1998). The Influence of Rhythmic Context on Schwa Epenthesis and Schwa Deletion in Dutch. *Language and Speech*, *41* (1), 87–108.

Lea, W. (1979). Testing linguistic stress rules with listeners' perception. In J. Wolf & D. H. Klatt (Eds.), *Speech Communication Papers presented at the 97th meeting of the ASA*. New York.

Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, *8*, 249–336.

Nespor, M., & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris Publications.

Prince, A. (1983). Relating to the grid. *Linguistic Inquiry*, *14*, 19–100.

Quené, H., & Kager, R. (1993). Prosodic sentence analysis without exhaustive parsing. In V. J. J. P. van Heuven & L. C. W. Pols (Eds.), *Analysis and Synthesis of speech: Strategic research towards high-quality TTS generation* (pp. 115–130). Mouton de Gruyter.

Selkirk, E. (1980). Prosodic domains in phonology: Sanskrit revisited. In M. Aronoff & M.-L. Kean (Eds.), *Juncture* (*Studia linguistica et philologica, 7*) (pp. 107–129). Saratoga, CA: Anma Libri.

Selkirk, E. (1986). On Derived Domains in Sentence Phonology. *Phonology Yearbook*, *3*, 371–405.

Streefkerk, B. M. (1997). Acoustical correlates of prominence: A design for research. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam, 21*, 131–142.

Streefkerk, B. M. (2002). Prominence. Acoustic and lexical/syntactic correlates. Dissertation ACLC. LOT Utrecht.

Streefkerk, B. M., Pols, L. C. W., & Ten Bosch, L. F. M. (1997). Prominence in read aloud sentences, as marked by listeners and classified automatically. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam, 21*, 101–116.

Streefkerk, B. M., & Pols, L. C. W. (1998). Prominence in read aloud Dutch sentences as marked by naive listeners. *Tagungsband KONVENS-98* (pp. 201–205). Frankfurt a.M.

O'Shaughnessy, D. (1976). Modelling Fundamental Frequency and its Relationship to Syntax, Semantics and Phonetics. Dissertation MIT, Cambridge, Mass.

O'Shaughnessy, D., & Allen, J. (1983). Linguistic modality effects on fundamental frequency in speech. *Journal of the Acoustic Society of America*, *74* (4), 1155–1171.

Visch, E. (1989). *A Metrical Theory of Rhythmic Stress Phenomena*. Dordrecht: Foris.

Phonetic variation or phonological difference?

The case of the early versus the late accent-lending fall in Dutch

Johanneke Caspers
Leiden University

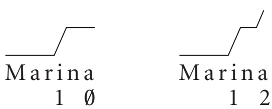
1. Introduction

Researchers investigating intonation generally work with a specific intonation model, since the variation occurring in natural speech data is so large that some abstract structure is called for. For Dutch there are two well-known models available: the phonetically oriented model developed at the Institute for Perception Research (IPO) in Eindhoven ('The Grammar of Dutch Intonation' or GDI, cf. 't Hart & Collier 1975; 't Hart, Collier, & Cohen 1990), and the phonologically inspired autosegmental model formulated by Gussenhoven (1984, 1988, 1991). In spite of the generally used qualifications 'phonetic' and 'phonological' respectively, *both* models constitute a "'phonological' but none the less phonetically accountable approach to intonation" (Ladd 1996: 42). Among other things, this means that both models are essentially abstract in nature. However, they do differ with respect to the definition of the relevant shapes of Dutch speech melody, and the present paper focuses upon one particular difference.

As an experimental linguist working on the meaning of melodic shapes in Dutch, I have an important question to answer: when is a melodic difference linguistically relevant, i.e., phonological in nature? Obviously, the notion of 'minimal pair' cannot be used to distinguish one melodic entity from another, since there is no list of 'melodic words' agreed upon. Also, melodic differences are seldom truth-functional in nature, especially the type of distinctions I am interested in, which means that a formal semantic approach cannot be used either.

Generally, the literature on intonation suggests consensus about the particular melodic shape ('phonetic interpretation') that is to be associated with a certain ('phonological') unit within a specific intonation model, and, vice versa, about the correct way to attach phonological labels to phonetic data. In practice, however, there does not appear to be such a clear relationship between phonological entities and phonetic data. The problem is not limited to rare melodic shapes; in fact, it is not even clear how the notion of *starred tone* – a fundamental part of all autosegmental models of intonation – has to be interpreted phonetically (Arvaniti, Ladd, & Mennen 2000).

As an illustration of the problem I give an example from my own experience. For an experiment meant to evaluate the meaning differences between two Dutch melodic shapes I used the data of two trained native speakers. They were instructed to read aloud different proper names with two different contour types, which were defined for them in terms of the Grammar of Dutch Intonation, and with drawings:



Both contours include an accent-lending rise ('1' in terms of the GDI) on the second – stressed – syllable; the first contour ends in level high pitch on the final syllable (indicated with the symbol 'Ø'), whereas the second one ends in a final rise ('2'). In a paper reporting the experiment, figures showing the pitch curves and segmental structure of all utterances produced by the two speakers were given to clarify the phonetic details. The melodic shapes were presented in GDI notation and they were translated into the autosegmental model as well; in consultation with Gussenhoven, the first melodic shape was labeled L*H (one of the three tone morphemes distinguished in his autosegmental analysis of Dutch intonation, the 'rise') and the second shape was labeled L*H H% (a 'rise' followed by a high boundary tone, H%, associated with the end of the domain, cf. Gussenhoven 1991; Van den Berg, Gussenhoven, & Rietveld 1992).

In different comments on the paper, three alternative autosegmental labelings for the two contours were suggested: (i) H*H L% versus H*H H%, (ii) H* versus L*H and (iii) half-completed L*H versus L*H (see Caspers 1998b). Remarkably, no arguments were given as to why a certain alternative autosegmental notation should be better than the one presented in the paper. There were no references to the available phonetic details. The alternative autosegmental transcriptions of the two contours under investigation were simply posed, indicating that there are problems with the concreteness of intonation models.

How does one decide which model gives the better interpretation of certain melodic phenomena? The present paper attempts an experimental approach to the question of how a specific melodic distinction in Dutch should be modelled phonologically.

Gussenhoven's model of Dutch intonation distinguishes – at least – two types of accent-lending fall, a *downstepped* and a *non-downstepped* fall (H*L versus !H*L), whereas the GDI recognizes only one type of accent-lending fall ('A'). In the autosegmental approach the difference between a downstepped fall and a non-downstepped one – when preceded by high level pitch – amounts to a difference in position relative to the accented syllable (cf. Rietveld & Gussenhoven 1995: Figure 1). In the implementation of the GDI for Dutch text-to-speech synthesis (Collier 1991) the precise timing of the accent-lending fall ('A') with respect to the syllable depends on the melodic environment it appears in. This means that the autosegmental model recognizes two types of falls, differing in timing with respect to the accented syllable, whereas the GDI considers these differently timed falls as phonetic variants of one underlying fall.

Can we decide which model gives the better analysis, i.e., whether differently timed pitch falls belong to separate basic categories or whether they are phonetic variants of a single phonological form? Or, to use Ladd's (1996) terminology, is the difference between an early and a late accent-lending fall in Dutch linguistic or paralinguistic?

2. Approach

What constitutes proof of a phonological difference between melodic forms? Within an experimental linguistic setting, two different approaches to this problem come to mind:

(1) categorical versus gradual perception

A widely accepted view is that a phonological difference should be categorical in nature, rather than scalar (e.g. Pierrehumbert & Steele 1989; Ladd 1996). There is a certain experimental technique for determining whether a difference between two shapes is perceived as gradual or categorical. This research paradigm was originally developed for segmental distinctions, but, recently, the same types of experiments were set up to investigate the nature of melodic differences. However, these categorical perception experiments do not always lead to conclusive results, as can be derived from the work by Ladd and Morton (1997) and Remijsen and van Heuven (this volume). Also, it may not even be the case that all phonological contrasts have to be strictly categorical in nature. In fact, Gerrits and Schouten (1998) could not find evidence in terms of strictly categorical perception for the unquestionable phonological differences between Dutch vowels.

A second possible approach is the following:

(2) differences in abstract meaning

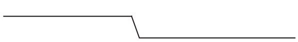
There seems to be – implicit – agreement that when two melodic shapes differ in meaning they have to differ phonologically (e.g. Pierrehumbert & Hirschberg 1990; Rietveld & Gussenhoven 1995; Ladd 1996; Grabe, Gussenhoven, Haan, Marsi, & Post 1998). Starting from this point, we may have an instrument to decide whether two melodic shapes are phonetic variants of the same underlying phonological form or whether they constitute two separate phonological entities.

The aim of the experiments presented below is to find out whether Dutch has two types of falling pitch accent or not by investigating whether there is a systematic meaning difference associated with the timing difference between an early and a late fall. To this end, an early and a late accent-lending fall are contrasted with each other and with the most frequently used pitch accent type in Dutch: the 'pointed hat', a rise immediately followed by a fall ('1&A' or H*L). This 'default' pitch accent type will function as a point of reference. Below the three pitch accent types under investigation are illustrated with a schematic contour:

(1)


(0)	1&A	(0)	(GDI)
%L	H*L	L%	(Gussenhoven)

Ik heb Jolanda gezien

(2)


(Ø)	A	(0)
%H	!H*L	L%

Ik heb Jolanda gezien

(3)

(Ø) &A (0)
%H H*L L%
Ik heb Jolanda gezien
I have Jolanda seen
'I have seen Jolanda'.

Throughout the remainder of the text, label '1&A' is used for the pointed hat; for the early fall label 'A' is used, and for the late fall '&A', because this movement is phonetically very akin to the fall in the pointed hat. Note that the early accent-lending fall ('A') and the late fall ('&A') only differ from each other in the timing domain. The late fall and the pointed hat ('1&A') differ in onset height: contour (3) starts with high pitch (indicated with the symbols 'Ø' or '%H') and contour (1) with low pitch ('0' or '%L'). Finally, 'A' and '1&A' differ in both onset height and timing of the fall.

3. Background and hypotheses

Two important semantic notions that have been associated with intonation are *prominence* and *phrasing*: speakers use melodic means to direct the listener's attention to the semantically central aspects of the message (prominence) and to aid the listener in segmenting the speech stream at several linguistic levels (phrasing). For the signalling of semantically central or 'focused' constituents, pitch accents are exploited, and the location of boundaries may be marked by specific boundary-marking pitch configurations. But why are there different *types* of pitch accent and different *types* of melodic boundary markers?

It is fairly commonly acknowledged that speech melody carries another semantic aspect in addition to the marking of prominence and phrasing: the signalling of information status with respect to the discourse (e.g. Gussenhoven 1984; Pierrehumbert & Hirschberg 1990). For example, all accent-lending pitch configurations carry the meaning of *prominence* ('this is important information'), but the type of pitch accent adds a further meaning component, which pertains to the status of the focused information with respect to a 'background' shared between speaker and listener.

Earlier experimental investigations provided support for the hypothesis (largely based on the theoretical analysis of Dutch melodic shapes formulated by Keijsper 1984) that a pointed hat ('1&A') marks information as *new*, whereas the early accent-lending fall ('A') is not very suitable to focus new information, but fits information that was *projected before the moment of speaking* (Caspers

1997, 1998a, 2000; Caspers, van Heuven, & van Zwol 1998). Furthermore, the pointed hat sounds polite and friendly while the late accent-lending fall ('&A') does not, and '&A' sounds detached and irritated whereas '1&A' does not (Grabe et al. 1998). This short summary of experimental findings indicates that there may be differences in meaning and attitude associated with the three melodic shapes defined above, but there were no direct comparisons between the early ('A') and the late ('&A') accent-lending fall.

On the basis of these experimental findings plus my own intuition, the following hypothesis was formulated:

– a difference in timing of a pitch fall corresponds to a difference between *new* and *projected* information

It is expected that the early fall 'A' will suit already projected information, whereas the late fall '&A' – as '1&A' – will fit new information.

Inspired by Judith Haan (p.c.) and Grabe et al. (1998: 65) it was further hypothesized that:

– a difference in onset height of a contour corresponds to a difference between *elliptic* and *non-elliptic* information

This means that a contour starting with a high onset (e.g. '&A') is expected to refer back to earlier information and therefore fit 'elliptic' information, whereas the low starting '1&A' does not. Formulated differently, elliptic utterances are expected to justify a high onset.

Note that the hypothesized differences between the three pitch contours are relatively small: they supposedly differ in expressed information status, but they share the primary meaning aspect of 'prominence'.

4. Experimental setup

The approach chosen to test the hypotheses formulated was a perception experiment, in which stimulus materials were presented to subjects in the form of a series of short conversations between two teachers (A and B) working at the same school, discussing the upcoming school party. Subjects were instructed to identify with speaker B.

The opposition between 'projected' and 'new' information was operationalized as an opposition between an utterance which forms the second part of a conjunction (B2) and the same utterance as a single 'new' utterance

(B1); the target utterance, carrying the melodic shape under investigation, is in italics; the first part of the conjunction is underlined:

A Er is nog niemand van de barcommissie.
There is yet nobody from the bar.comittee
'There is nobody from the bar committee yet'.

B 1 *Ik heb Jolanda gezien.*
I have Jolanda seen
'I have seen Jolanda'.

B 2 Ik hoorde Marina en *ik heb Jolanda gezien.*
I heard Marina and I have Jolanda seen
'I heard Marina and I have seen Jolanda'.

The prediction was that the early accent-lending fall ('A') will not be very suitable when the focused information is new (B1), whereas it would fit the already projected information given in the second half of a conjunction (B2).

The opposition between 'plus ellipsis' and 'minus ellipsis' was operationalized as the difference between utterances implicitly (B1) and explicitly (B2) contradicting the previous utterance (i.e., the one produced by speaker A); the explicit refutation of A's statement is in bold face:

A Er is nog niemand van de barcommissie.
There is yet nobody from the bar.committee
'There is nobody from the bar committee yet'.

B 1 *Ik heb Jolanda gezien.*
I have Jolanda seen
'I have seen Jolanda'.

B 2 **Jawel!** *Ik heb Jolanda gezien.*
Yes.there.is! I have Jolanda seen
'Yes there is! I have seen Jolanda'.

In B1, the reply to A's utterance can be viewed as 'elliptic', because it implicitly rejects A's statement; according to the hypothesis formulated above, a contour with a high onset would therefore fit best (e.g. '&A'). In B2, speaker B starts his reply with an explicit rejection of A's utterance (bold faced 'Jawel!'), which, according to the second hypothesis, would allow a low onset ('1&A').

The opposition between 'plus' and 'minus ellipsis' is not considered relevant for the utterances containing 'projected' information, because it seems impossible to have an utterance that is projected by an earlier utterance (e.g. the first part of a conjunction) and at the same time elliptic (i.e., the preceding utterance has been omitted).

Summarizing, the subjects were presented with either:

1. A + *target*
2. A + **negation** + *target*

or

3. A + **negation** + conjunction Part 1 + *target*

As indicated above, the early accent-lending fall 'A' was expected to fit 'projected' information (3) but not 'new' information (1 and 2). On the basis of earlier research it was further expected that the default pitch accent ('1&A') fits 'projected' information as well as 'new' information, since it seems possible to interpret virtually all focused information as somehow 'new' (cf. Caspers 2000; Caspers et al. 1998); however, '1&A' is not expected to fit 'elliptic' information (1). Since the late fall ('&A') supposedly marks focused information as 'new' and 'elliptic', this shape is expected to fit context type (1); '&A' may also be acceptable in context type (3), because the 'elliptic'-'non-elliptic' opposition seems only relevant at the onset of a contour.

In addition to the example given above, three further sets of stimuli were included in the design. The target utterances all consist of seven syllables, with an accented middle syllable, which is also the middle syllable of a proper name. Care was taken that the target utterances themselves were as neutral as possible with respect to attitude, which means that e.g. intrinsically reproachful utterances such as "Je was gestopt met roken!" ('But you quit smoking!') were avoided. Furthermore, utterances that could easily be interpreted as sensational, like "Hij is boeddhist geworden" ('He turned buddhist') were avoided, because these would preclude the use of the early fall 'A' (cf. Caspers et al. 1998).

A Jan-Willem heeft nog niks voor de muziek geregeld.
Jan-Willem has yet nothing for the music organized
'Jan-Willem hasn't organized anything for the music yet'.

B 1 *Hij heeft Cornelis gehaald.*
He has Cornelis fetched
'He has fetched Cornelis'.

B 2 **Jawel!** *Hij heeft Cornelis gehaald.*
Yes.he.did! He has Cornelis fetched
'Yes he did! He has fetched Cornelis'.

B 3 **Jawel!** Hij heeft alles op het podium gezet en *hij heeft*
Yes.he.did! He has everything on the stage put and he has
Cornelis gehaald.
Cornelis fetched
'Yes he did! He has put everything on stage and he has fetched Cornelis'.

A Jolanda heeft nog niemand benaderd voor de opruimploeg.
Jolanda has yet nobody approached for the cleaning.mob
'Jolanda hasn't approached anyone for the cleaning mob yet'.

B 1 *Ze heeft Marina gebeld*.
She has Marina called
'She has called Marina'.

B 2 **Jawel!** *Ze heeft Marina gebeld*.
Yes.she.did! She has Marina called
'Yes she did! She has called Marina'.

B 3 **Jawel!** Ze heeft Jan-Willem gevraagd en *ze heeft Marina*
Yes.she.did! She has Jan-Willem asked and she has Marina
gebeld.
called
'Yes she did! She has asked Jan-Willem and she has called Marina'.

A We kunnen niet in de sportzaal terecht.
We can not in the gym be.admitted
'We cannot use the gym'.

B 1 *Ik heb Jan-Willem bepraat*.
I have Jan-Willem persuaded
'I have persuaded Jan-Willem'.

B 2 **Jawel!** *Ik heb Jan-Willem bepraat*.
Yes.we.can! I have Jan-Willem persuaded
'Yes we can! I have persuaded Jan-Willem'.

B 3 **Jawel!** Ik heb een losse vloer besteld en *ik heb*
Yes.we.can! I have a detachable floor ordered and I have
Jan-Willem bepraat.
Jan-Willem persuaded
'Yes we can! I have ordered a detachable floor and I have persuaded Jan-Willem'.

4.1 Design

The investigation comprised two types of test: a pairwise comparison experiment and a rating experiment. The first experiment was designed to test the hypotheses formulated in §3. The goal of the second experiment was twofold: providing additional information to the pairwise comparison data, and verifying whether there is a difference in attitude associated with the three pitch accent types.

In the comparison experiment subjects had to compare two melodic versions of a target utterance in a specific dialogue context, and they had to select the contour best fitting the presented context. In the rating experiment subjects were asked to judge each combination of context and contour type (using the same materials) on the following scales: acceptability, irritation, detachment and finality. 'Acceptability' is used as a complement to the pairwise comparison data; 'irritation' and 'detachment' are two of the 'favorability' scales used by Grabe et al. (1998), and 'finality' is an attitude associated with downstep (cf. Rietveld & Gussenhoven 1995; Swerts, Bouwhuis, & Collier 1994; Ladd 1996).

4.2 Method

Two Dutch intonologists, a male and a female, realized the four target utterances with each of the three intonation contours. Figures representing the pitch curves of the stimuli can be found in the appendix. For clarification, the natural pitch contours can be compared with the schematic contours given in §1.

Thirty-six native Dutch listeners participated in the experiments.[1] Their ages varied between 21 and 58; no hearing difficulties were reported and they were paid. The data were presented via an interactive computer program.[2] Subjects needed approximately 30 minutes to complete the task. Since the majority of subjects participated through Internet, there was no strict control over the circumstances under which the experiment was performed (such as ambient noise, type of headphones, type of loudspeaker, etc.).

To avoid a direct influence of the pitch of the context preceding the target utterance, subjects were presented with only visual representations of the dialogue contexts.

In the pairwise comparison experiment subjects were asked to picture themselves as speaker B in each of the (visually) presented dialogue contexts and to decide which of **two** melodic versions of the target utterance – which they could make audible as often as they wished – best fitted the given situation, indicating the confidence of their choice (sure/unsure).

In the rating experiment subjects were asked to judge the combination of a specific dialogue context and the melody of a specific (audible) target utterance on a ten-point scale, ranging from e.g. totally unacceptable intonation ('1') to totally acceptable intonation ('10'). In the Dutch educational system the values 1 to 5 represent degrees of inadequacy, whereas values 6 to 10 represent degrees of adequacy; the boundary between acceptable and unacceptable is drawn at 5.5.

The order of the two experiments as well as the order of the four scales within the rating experiment was counter-balanced over subjects.

5. Results

5.1 Pairwise comparison experiment

Table 1 presents the preference judgements for the three contour types, broken down by the three context types (note that the maximum percentage per contour type is 67%, since each of the three contour types is present in only two-thirds of the pairwise comparisons, because subjects had to choose between two contours, not between three).

The strongest context effect lies in the opposition between 'new' and 'projected' information; in both 'new' context types, the pointed hat ('1&A') and the late fall ('&A') are preferred over the early fall ($\chi^2 = 62.87$, df = 2, p < .001; $\chi^2 = 70.34$, df = 2, p < .001), whereas there is no preference for any contour type in the 'projected' contexts ($\chi^2 = .721$, df = 2, ins.). There is no significant effect of the factor 'ellipsis' ($\chi^2 = 2.01$, df = 2, ins.), which means that there does not seem to be a systematic relationship between the height of the onset of the stimulus utterance and whether or not the utterance can be viewed as elliptic.

Table 1. Absolute frequency (and percentage) of preferred contour type, broken down by context type

	context type			
preferred contour type	**new + ellipsis**	**new – ellipsis**	**projected**	**total**
'A'	61 (15%)	62 (15%)	131 (32%)	254 (21%)
'&A'	166 (41%)	147 (36%)	133 (33%)	446 (36%)
'1&A'	181 (44%)	199 (49%)	144 (35%)	524 (43%)
total	408 (100%)	408 (100%)	408 (100%)	1224 (100%)

The early accent-lending fall ('A') differs from the late fall ('&A') and the pointed hat ('1&A') in that it does not fit the 'new' contexts very well (preferred in only 15% of the cases), which can be taken as support for the hypothesis that the early and the late accent-lending fall form different phonological categories.

Considering the confidence of the subjects when making their choices (Table 2), it becomes clear that it is easier to make a choice between a pointed hat and an early accent-lending fall ('A' versus '1&A'), than between a late fall and an early fall or between a late fall and a pointed hat (χ^2 = 14.67, df = 2, $p < .001$). This finding is completely in line with the earlier observed formal differences between the three contour types: 'A' differs form '1&A' in both onset height and timing of the fall, whereas the other two oppositions involve just one dimension.

Table 2. Absolute frequency (and percentage) of sure and unsure preference responses, broken down by choice of contours

	choice of contours			
confidence	**'A' versus '1&A'**	**'A' versus '&A'**	**'&A' versus '1&A'**	**total**
unsure	112 (27%)	159 (39%)	155 (38%)	426 (35%)
sure	296 (73%)	249 (61%)	253 (62%)	798 (65%)

Table 3 presents the preferences for the three contour types, separately for the four stimulus sentences.

Table 3. Absolute frequency (and percentage) of preferred contour type, broken down by stimulus sentence

	stimulus sentence			
preferred contour type	**'Cornelis'**	**'Jan-Willem'**	**'Jolanda'**	**'Marina'**
'A'	57 (18%)	69 (22%)	70 (23%)	58 (19%)
'&A'	109 (36%)	109 (36%)	112 (37%)	116 (38%)
'1&A'	140 (46%)	128 (42%)	124 (40%)	132 (43%)

There is no significant difference between the preference scores for the four sentence types (χ^2 = 3.57, df = 6, ins.).

Finally, Table 4 gives the scores per contour type, separately for the two speakers.

The table shows that the pointed hat contour ('1&A') is preferred more often when spoken by the male speaker, whereas the early accent-lending fall ('A') is preferred more often when realized by the female speaker (χ^2 = 12.03, df = 2,

Table 4. Absolute frequency (and percentage) of preferred contour type, broken down by speaker

preferred contour type	female speaker	male speaker	total
'A'	150 (25%)	104 (17%)	254 (21%)
'&A'	222 (36%)	224 (37%)	446 (36%)
'1&A'	240 (39%)	284 (46%)	524 (43%)

p < .01). For the late fall, the preference scores for the two speakers are virtually identical. There is no obvious explanation for these (small) speaker effects.

5.2 Rating experiment

5.2.1 *Acceptability scores*

In Table 5 the mean acceptability scores are given for the three contour types, broken down by the three context types.

Table 5. Mean acceptability scores (and standard deviation) per contour type, broken down by context type

	context type			
contour type	new + ellipsis	new – ellipsis	projected	mean
'A'	5.1 (2.1)	6.5 (2.3)	6.5 (2.5)	6.0 (2.4)
'&A'	6.2 (2.4)	7.4 (1.7)	6.4 (2.2)	6.7 (2.2)
'1&A'	7.3 (2.0)	8.0 (2.3)	7.3 (2.3)	7.5 (2.2)
mean	6.2 (2.3)	7.3 (2.2)	6.7 (2.3)	6.7 (2.3)

Overall, the stimuli were judged as fairly acceptable (a mean score of 6.7). An analysis of variance ($\alpha = .05$) shows that there is a significant effect of contour type ($F_{(2,303)} = 11.6$, $p < .001$) and an effect of context type ($F_{(2,303)} = 6.7$, $p < .005$), but no significant interaction between these factors ($F_{(4,297)} = 1.1$, ins.). This means that the influence of pitch accent type is approximately the same for the three context types, and that the effect of context type is the same for the three pitch accent types. Post-hoc analyses (Newman-Keuls) reveal that the pointed hat is judged significantly more acceptable (a mean score of 7.5) than the late accent-lending fall (6.7), which in turn sounds significantly more acceptable than the early accent-lending fall (6.0). Furthermore, the stimuli in the 'new minus ellipsis' contexts are considered more acceptable (a mean score of 7.3) than those in the 'new plus ellipsis' contexts (6.2, an effect that may be attributable to the 'abrupt' nature of the elliptic stimuli), but they both do not differ statistically from the 'projected' contexts (6.7).

The acceptability ratings confirm the tendencies found in the pairwise comparison test, in the sense that the pointed hat is judged as most acceptable; however, the fact that the pointed hat ('1&A') and the late fall ('&A') were preferred over the early fall ('A') in the 'new' contexts is not reflected in significantly lower acceptability scores for 'A' than for '&A' and '1&A' in the 'new plus' and 'new minus' conditions. Note, however, that the only 'insufficient' mark (5.1) can be found for 'A' in the 'new plus ellipsis' condition, suggesting that an early fall as initial pitch accent on a 'new' utterance is not very acceptable.

5.2.2 *Irritation scores*

Table 6 contains the mean irritation scores for the relevant variables.

Table 6. Mean irritation scores (and standard deviation) per contour type, broken down by context type

	context type			
contour type	**new + ellipsis**	**new – ellipsis**	**projected**	**mean**
'A'	4.8 (2.3)	4.6 (2.3)	5.9 (2.5)	5.1 (2.4)
'&A'	5.1 (2.7)	4.5 (2.6)	4.4 (2.4)	4.7 (2.6)
'1&A'	3.6 (2.4)	4.4 (2.3)	2.7 (1.6)	3.6 (2.2)
mean	4.5 (2.5)	4.5 (2.4)	4.3 (2.6)	4.4 (2.5)

First of all, the stimuli did not sound very irritated at all: an overall mean score of 4.4, which is well below the critical value of 5.5. An analysis of variance on the irritation judgments shows an effect of contour type ($F_{(2,302)} = 11.1$, $p < .001$), no main effect of context type ($F_{(2,302)} < 1$), and a clear interaction between contour and context type ($F_{(4,296)} = 4.2$, $p < .005$). The early fall ('A') on a 'projected' context sounds more irritated (5.9) than the late fall (4.4) or the pointed hat (2.7), whereas the three contours receive practically equal scores in the 'new minus ellipsis' context (4.6, 4.5 and 4.4 respectively). These impressions are confirmed by post-hoc analyses: when the context contains projected information 'A' sounds significantly more irritated than '&A', which in turn sounds more irritated than '1&A'; in the 'new minus ellipsis' contexts there is no effect of contour type, and in the 'new plus ellipsis' contexts the late fall sounds significantly more irritated than the pointed hat.

These data do not suggest a clear-cut difference in 'irritation' expressed by the three types of pitch accent; however, there does seem to be some systematic effect of contour type under specific contextual circumstances.

5.2.3 *Detachment scores*

Table 7 presents the mean detachment scores, which all remain below 6. This indicates that the stimuli did not sound very detached.

Table 7. Mean detachment scores (and standard deviation) per contour type, broken down by context type

	context types			
contour type	**new + ellipsis**	**new – ellipsis**	**projected**	**mean**
'A'	5.6 (2.1)	5.6 (2.8)	5.5 (2.0)	5.6 (2.3)
'&A'	4.4 (2.2)	4.5 (2.2)	4.1 (2.1)	4.3 (2.2)
'1&A'	4.6 (2.5)	4.3 (2.4)	3.9 (2.1)	4.3 (2.3)
mean	4.9 (2.3)	4.8 (2.5)	4.5 (2.2)	4.7 (2.3)

An analysis of variance on the detachment scores reveals a main effect of contour type ($F_{(2,303)} = 12.3$, $p < .001$), but no influence of context ($F_{(2,303)} = 1.1$, ins) and no interaction ($F_{(4,297)} < 1$). Post-hoc analyses show that 'A' sounds more detached than '1&A' or '&A'. In addition – and in contrast with the results found for the other scales – there is a rather distinct influence of speaker on the detachment scores ($F_{(1,304)} = 29.7$, $p < .001$), as well as small interaction effects between speaker and context type ($F_{(2,299)} = 6.7$, $p < .005$) and between speaker and contour type ($F_{(2,299)} = 3.1$, $p < .05$):

Table 8. Mean detachment scores (and standard deviation) per contour type and per context type, broken down by speaker

		context type		
contour type	**speaker**	**new + ellipsis**	**new – ellipsis**	**projected**
'A'	female	5.2 (1.8)	3.6 (2.2)	5.3 (2.3)
	male	6.2 (2.2)	7.4 (1.8)	5.6 (1.6)
'&A'	female	4.2 (2.5)	3.9 (2.3)	4.3 (2.3)
	male	4.8 (1.9)	5.1 (2.0)	3.9 (2.0)
'1&A'	female	4.1 (2.1)	2.9 (1.4)	3.1 (1.7)
	male	5.2 (2.8)	5.6 (2.4)	4.6 (2.1)

The difference between the two speakers is largest in the 'new minus ellipsis' contexts (a mean detachment score of 6.0 for the male speaker versus 3.5 for the female speaker), and smallest for the late falls (4.6 versus 4.1). In all but one case (i.e., the late fall in the 'projected' contexts) the male speaker sounds more detached than the female speaker.

These results indicate that, apart from possible additional speaker influences, 'detachment' is an attitude closely associated with the early accent-lending fall.

5.2.4 *Finality scores*

Table 9 gives the finality scores.

Table 9. Mean finality scores (and standard deviation) per contour type, broken down by context type

contour types	context type new + ellipsis	 new – ellipsis	 projected	mean
'A'	6.6 (2.5)	6.8 (2.4)	7.6 (2.4)	7.0 (2.5)
'&A'	6.1 (2.6)	6.3 (2.7)	6.5 (2.3)	6.3 (2.5)
'1&A'	7.1 (2.3)	8.2 (2.0)	7.3 (2.2)	7.5 (2.2)
mean	6.6 (2.5)	7.1 (2.5)	7.1 (2.3)	6.9 (2.4)

The finality scores show an effect of contour type ($F_{(2,300)} = 6.6$, $p < .005$), no difference between the three context types ($F_{(2,300)} = 1.7$, ins.), and no interaction ($F_{(4,294)} = 1.2$, ins.). All contour types sound rather final, but post-hoc analyses reveal that the late fall sounds significantly less final (a mean score of 6.3) than the early fall or the pointed hat (7.0 and 7.5 respectively). This means that there is support for the hypothesis that the early (downstepped) fall ('A') sounds more final than the late fall ('&A').

6. Conclusion

Given the fact that the subjects could not be forced to interpret stimuli occurring as the second part of a conjunction (context type 3) as 'projected information' (cf. §4), plus the fact that there is a large overlap in meaning between the three pitch accent types (viz., 'plus focus' or "this is important information"), plus the fact that one cannot be completely sure that subjects really did what was asked of them ("identify with speaker B"), the results of the experiments seem quite clear.

The results from the preference test and the acceptability ratings provide support for the hypothesis that a difference in timing corresponds to a difference between 'new' and 'projected' information: the early accent-lending fall 'A' does not fit new information very well, while it suits projected information, whereas both the late fall ('&A') and the pointed hat ('1&A') fit information that can be interpreted as new. This means that the predictions with regard to the differences between the three melodic shapes in terms of information status (cf. §4) were borne out for the 'projected' – 'new' opposition. However, there was no support for the second hypothesis: '&A' was not preferred over '1&A' in the 'elliptic' contexts. This may be due to the way the stimulus material was set up: the elliptic version of B's utterance should have been interpreted as a rejection of A's statement, which may not have been clear to all subjects (they may have interpreted the B1 utterances as independent statements instead of reactions to A's utterances).

With respect to the main question – is there as systematic meaning difference between an early and a late fall in Dutch? – the results can roughly be summarized as follows:

– the early accent-lending fall ('A') does not fit new information, in contrast with the late accent-lending fall ('&A')
– the early accent-lending fall is perceived as less acceptable, more detached, more final and more irritated than the late accent-lending fall

This indicates that the two types of fall differ in information status as well as expressed attitude, which in my view should lead to the tentative conclusion that they must belong to two separate phonological categories: Dutch has two different types of accent-lending fall (supporting the analysis of Gussenhoven).[3] For a more firm statement, further research involving more speakers and more stimulus material is called for, followed by production studies (including database research) and imitation studies (cf. Pierrehumbert & Steele 1989).

As explained in the introduction, the linguistic relevance of melodic phenomena cannot be determined in the same way as, for instance, segmental phenomena (phonemes), because there is no list of melodic words that makes it possible to look for minimal pairs. This means that other criteria have to be used, and experimentally verified meaning differences seem fruitful for improving the available models of speech melody. Without experimental investigation of what speakers do and hearers perceive, the analysis of melodic phenomena remains merely introspective, which can be rather tricky, as I have tried to explain in the introduction to this paper.

Acknowledgment

This research was funded by the Netherlands Organization for Scientific Research (NWO), under project #300–75–001.

Notes

1. As a result of a computer crash and a mistake in interpreting the instructions, the data of two subjects could not be used in the analyses.

2. URL http://fonetiek-6.leidenuniv.nl/caspers/le5-intro.html, programmed by J. J. A. Pacilly B.Sc. of the Leiden University Phonetics Laboratory.

3. The Grammar of Dutch Intonation seems to be wrong, in the sense that it does not distinguish two types of accent-lending fall. However, the grammar does distinguish a special type of fall, preceded by a small high rise ('5&A'), which may be identical or at least very close to the late fall without such a rise ('&A').

References

Arvaniti, A., Ladd, D. R., & Mennen, I. (2000). What is a starred tone? Evidence from Greek. In M. Broe & J. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V* (pp. 119–131). Cambridge: Cambridge University Press.

Berg, R. van den, Gussenhoven, C., & Rietveld, T. (1992). Downstep in Dutch: Implications for a model. In G. J. Docherty & D. R. Ladd (Eds.), *Laboratory Phonology II* (pp. 335–359). Cambridge: Cambridge University Press.

Caspers, J. (1997). Testing the meaning of four Dutch pitch accent types. *Proceedings 5th European Conference on Speech Communication and Technology*, Vol. 2, 863–866.

Caspers, J. (1998a). Experiments on the meaning of two pitch accent types: The 'pointed hat' versus the accent-lending fall in Dutch. *Proceedings International Conference on Spoken Language Processing 1998*, 1291–1294.

Caspers, J. (1998b). Who's next? The melodic marking of question versus continuation in Dutch. *Language and Speech, 41*, 371–394.

Caspers, J. (2000). Experiments on the meaning of four types of single-accent intonation contours in Dutch. *Language and Speech, 43*, 127–161.

Caspers, J., van Heuven, V. J., & van Zwol, N. (1998). Experiments on the semantic contrast between the 'pointed hat' contour and the accent-lending fall in Dutch. In R. van Bezooijen & R. Kager (Eds.), *Linguistics in the Netherlands 1998* (pp. 65–79). Amsterdam: John Benjamins.

Collier, R. (1991). Multi-language intonation synthesis. *Journal of Phonetics, 19*, 61–73.

Gerrits, E., & Schouten, B. (1998). Categorical perception of vowels. *Proceedings International Conference on Spoken Language Processing 1998*, 2279–2282.

Grabe, E., Gussenhoven, C., Haan, J., Marsi, E., & Post, B. (1998). Preaccentual pitch and speaker-attitude in Dutch. *Language and Speech, 41*, 63–85.

Gussenhoven, C. (1984). *On the grammar and semantics of sentence accents*. Dordrecht: Foris.

Gussenhoven, C. (1988). Adequacy in intonation analysis: The case of Dutch'. In H. van der Hulst & N. Smith (Eds.), *Autosegmental studies on pitch accent* (pp. 95–121). Dordrecht: Foris

Gussenhoven, C. (1991). Tone segments in the intonation of Dutch. In T. F. Shannon & J. P. Snapper (Eds.), *The Berkeley Conference on Dutch Linguistics 1989* (pp. 139–155). Lanham, MD: University Press of America.

Hart, J. 't, & Collier, R. (1975). Integrating different levels of intonation analysis. *Journal of Phonetics, 3*, 235–255.

Hart, J. 't, Collier, R., & Cohen, A. (1990). *A perceptual study of intonation*. Cambridge: Cambridge University Press.

Keijsper, C. E. (1984). Vorm en betekenis in Nederlandse toonhoogtecontouren. *Forum der Letteren, 25*, 20–37, 113–126.

Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.

Ladd, D. R., & Morton, R. (1997). The perception of intonational emphasis: Continuous or categorical? *Journal of Phonetics, 25*, 313–342.

Remijsen, B., & van Heuven, V. J. (this volume). On the categoriality of intonational contrasts, an experiment on boundary tones in Dutch.

Pierrehumbert, J. B., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in communication* (pp. 271–311). Cambridge, MA: MIT Press.

Pierrehumbert, J. B., & Steele, S. A. (1989). Categories of tonal alignment. *Phonetica, 46*, 181–196.

Rietveld, T., & Gussenhoven, C. (1995). Aligning pitch targets in speech synthesis: effects of syllable structure. *Journal of Phonetics, 23*, 375–385.

Swerts, M., Bouwhuis, D. G., & Collier, R. (1994). Melodic cues to the perceived "finality" of utterances. *Journal of the Acoustical Society of America, 96*, 2064–2075.

Appendix: Stimulus contours

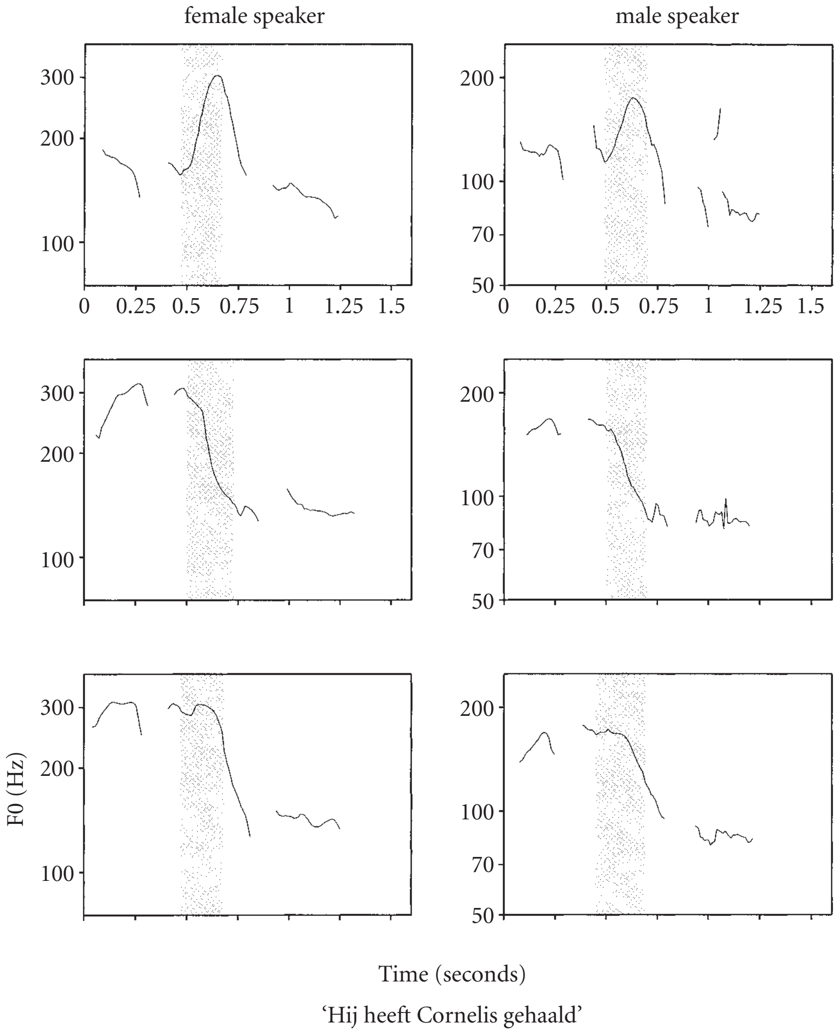


Figure 1. Stimulus pitch contours (in Herz) realized by the female and male speaker on the sentence "Hij heeft Cornelis gehaald"; grey rectangles mark accented syllables.

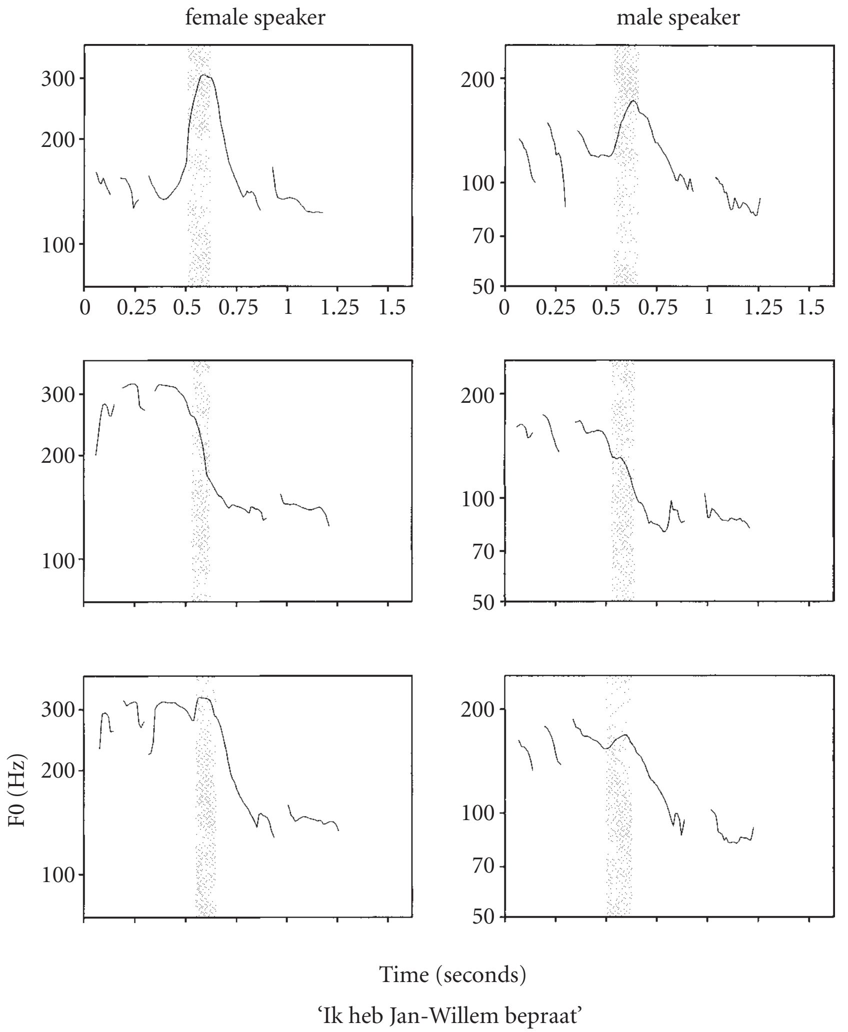


Figure 2. Stimulus pitch contours (in Herz) realized by the female and male speaker on the sentence "Ik heb Jan-Willem bepraat"; grey rectangles mark accented syllables.

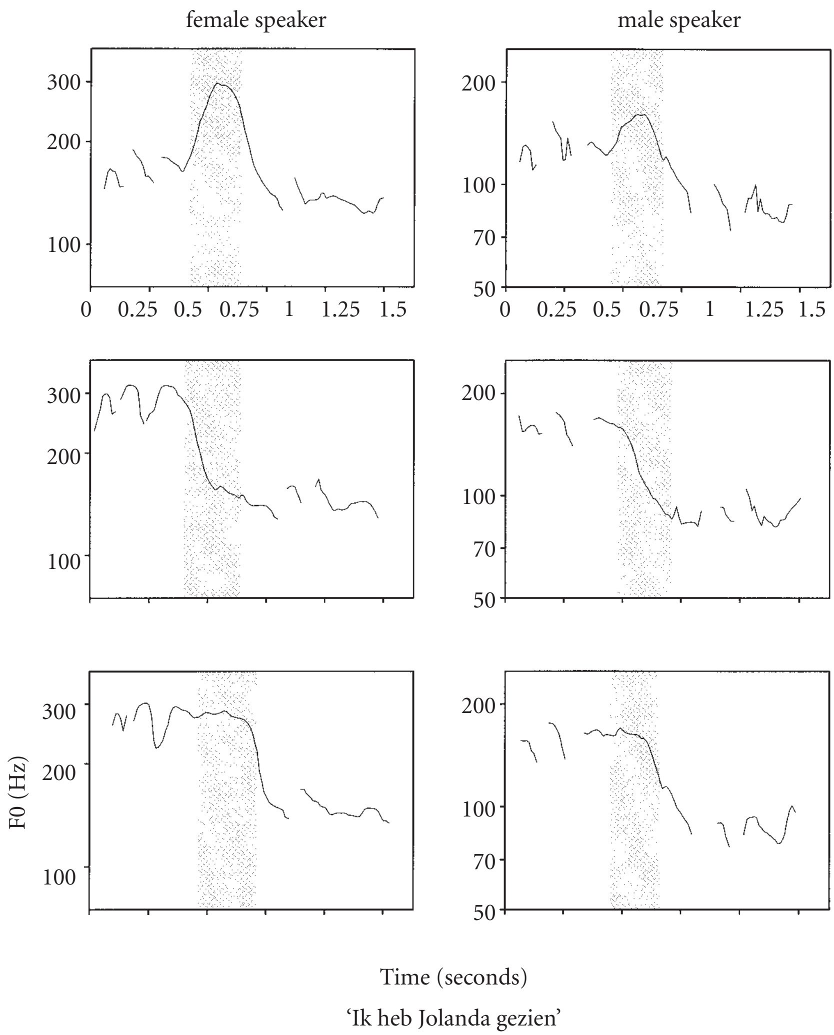


Figure 3. Stimulus pitch contours (in Herz) realized by the female and male speaker on the sentence "Ik heb Jolanda gezien"; grey rectangles mark accented syllables.

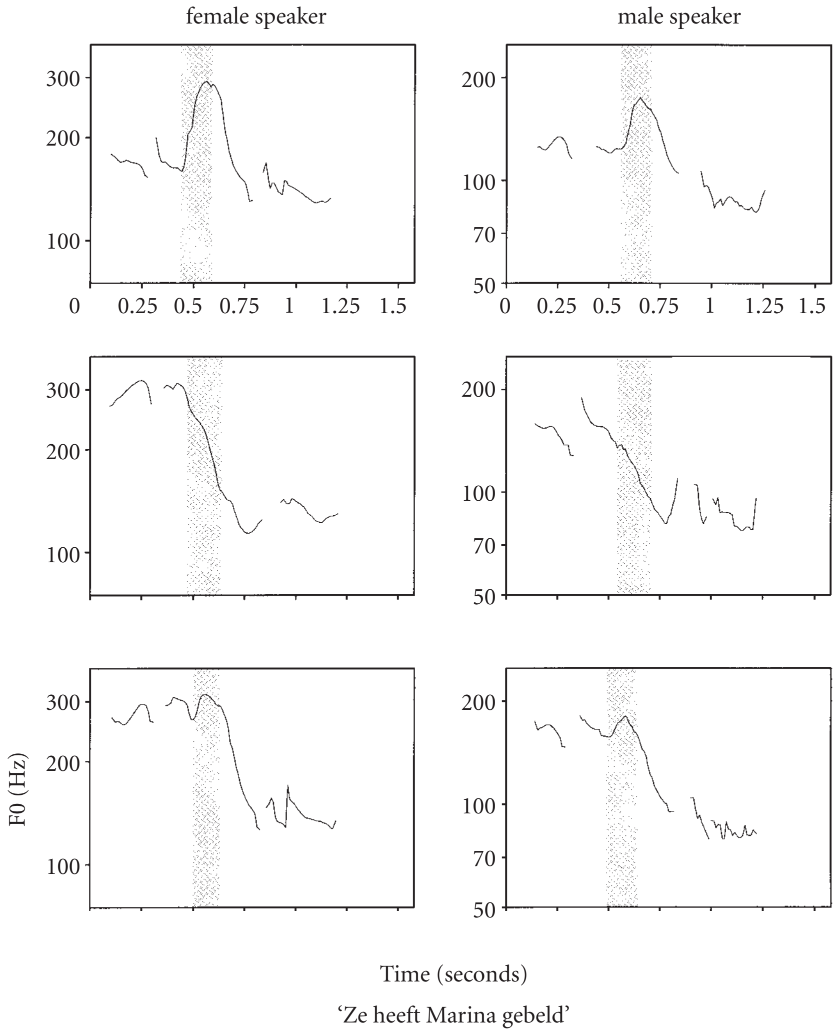


Figure 4. Stimulus pitch contours (in Herz) realized by the female and male speaker on the sentence "Ze heeft Marina gebeld"; grey rectangles mark accented syllables.

On the categorical nature of intonational contrasts

An experiment on boundary tones in Dutch*

Bert Remijsen and Vincent J. van Heuven
Leiden University

1. Introduction

1.1 Linguistic categorisation of sound

A basic problem of linguistic phonetics is to explain how the infinite variety of speech sounds in actual utterances can be described with finite means, such that they can be dealt with in the grammar, i.e. phonology, of a language. The crucial concept that was developed to cope with this reduction problem is the sound category, or – when applied to the description of segmental phenomena – the phoneme. This is best conceived of as an abstract category that contains all possible sounds that are mutually interchangeable in the context of a minimal word pair. That is, substitution of one token (allophone) of a phoneme for an other does not yield a different word (i.e., a string of sounds with a different lexical meaning).[1]

The phonemes in a language differ from one another along a finite number of phonetic dimensions, such as degree of voicing, degree of noisiness, degree of nasality, degree of openness, degree of backness, degree of rounding, etc. Each phonetic dimension, in turn, is subdivided into a small number (two to four) of phonologically functional categories, such as voiced/voiceless, (half)closed/(half)open, front/central/back, etc. Phonetic dimensions generally have multiple acoustical correlates. For instance, degree of voicing correlates with a multitude of acoustic cues such as voice onset time, duration of preceding vowel, steepness of intensity decay and of formant bends in preceding vowel, duration of intervocalic (near) silence, duration and intensity of noise burst, steepness of intensity attack and formant bends of following vowel. These acoustic properties typically covary in preferred patterns, but may be manipulated independently through speech synthesis. When non-typical ('conflicting') combinations of parameter values are generated in the laboratory, some cues prove to be more influential than others; so-called 'cue trading relationships' have been established for many phonemic contrasts. In Dutch, for instance, vowel quality (acoustically defined by F_1 and F_2, i.e., the centre frequencies of the lowest two resonances in the vocal tract) and vowel duration were found to be equally influential in cuing the tense/lax-contrast between /aː/ and /ɑ/: a duller vowel quality (lower F_1 and F_2-values), normally cuing /ɑ/ could be compensated for by increasing the duration of the vowel so that native listeners still perceive /aː/ (and vice versa, van Heuven 1986).

Categorisation of sounds may proceed along different lines. First, many differences between sounds are simply too small to be heard at all: these are subliminal. The scientific discipline of psycho-acoustics provides a huge literature on precisely what differences between sounds can and cannot be heard with the naked ear. Moreover, research has shown that the human hearing mechanism (and that of mammals in general) has developed specific sensitivities to certain differences between sounds and is relatively deaf to others. These predilections have been shown to be present at birth (probably even in utero), and need not be acquired through learning. However, human categorisation of sound is further shaped by exposure to language. As age progresses from infancy to adulthood, sound differences that were still above threshold shortly after birth, quickly lose their distinctivity. An important concept in this context is the notion of categorical perception. This notion is best explained procedurally in terms of a laboratory experiment.

Imagine a minimal word pair such as English *back* ∼ *pack*. One important difference between these two tokens is that the onset of voicing in *back* is more or less coincident with the plosive release, whilst the voice onset in *pack* does not start until some 50 ms after the release. It is not too difficult in the laboratory to create a series of exemplars by interpolating the voice onset time of a prototypical *back* (0 ms delay) and that of a prototypical *pack* (70 ms delay) in steps of, say, 10 ms, so that we now have an 8-step continuum ranging over 0, 10, 20, 30, 40, 50, 60, and 70 ms. These eight exemplars are shuffled in random order and played to an audience of native English listeners for identification as either *back* or *pack* (forced choice). The 0 ms voice delay token will naturally come out of the experiment with exclusively *back*-responses (0% *pack*); the 70 ms token will have 100% *pack*-responses. But what results will be obtained for the intermediate exemplars? If the 10 ms changes in voice delay are perceived *continuously*, one would predict a constant, gradual increase in %-*pack* responses for each 10 ms increment in the delay. I.e., when the stimulus increment (from left to right) is plotted against the response increment (from bottom to top), the psychometric function (the line that captures the stimulus-response relationship) is essentially a straight line (open symbols in Figure 1B). The typical outcome of experiments with voiced/voiceless continua, however, is non-continuous. For the first part of the continuum all exemplars are perceived as *back*-tokens, the rightmost two or three exemplars are near-unanimously perceived as *pack*. Only for one or two exemplars in the middle of the continuum do we observe uncertainty on the part of the listener: here the distribution of responses is more or less ambiguous between *back* and *pack*. The psychometric function for this so-called categorical perception is sigmoid, i.e., has the shape of an S (big solid symbols in Figure 1B). In the idealized case of perfect categorical perception we would, in fact, expect to see a step-function jumping abruptly from (almost) 0 to (almost) 100% *pack*-responses somewhere along the continuum (thin black line with small solid symbols in Figure 1B).

The *category boundary* (at 35 ms VOT in Figure 1B) is defined as the (interpolated) point along the stimulus axis where the distribution of responses is completely ambiguous, i.e., 50–50%. For a well-defined *cross-over* from one category to the other there should be a point along the stimulus axis where 75% of the responses agree on one category, and a second point where there is 75%-agreement on the other category. The *uncertainty margin* is defined in absolute terms as the distance along the stimulus axis between the two 75%-points; equivalent relative measures can be derived from the steepness of the psychometric function (e.g. the slope coefficient or the standard deviation of the cumulative normal distribution fitted to the data points).

Although a pronounced sigmoid function (such as the one drawn in Figure 1B) is a clear sign of categorical perception, researchers have always been reluctant to consider it definitive proof. Listeners, when forced to, tend to split any continuum down the middle. For a continuum to be perceived categorially, therefore, two conditions should be met:

– results of an identification experiment should show a clear sigmoid function, and
– the discrimination function should show a local peak for stimuli straddling the category boundary.

The discrimination function is determined in a separate experiment in which either (i) identical or (ii) adjacent tokens along the stimulus continuum are presented pairwise. Listeners then decide for each pair whether the two tokens are 'same' or 'different'. Two kinds of error may occur in a discrimination task:

– a physically different pair may be heard as same, and
– a pair of identical tokens may be called different.

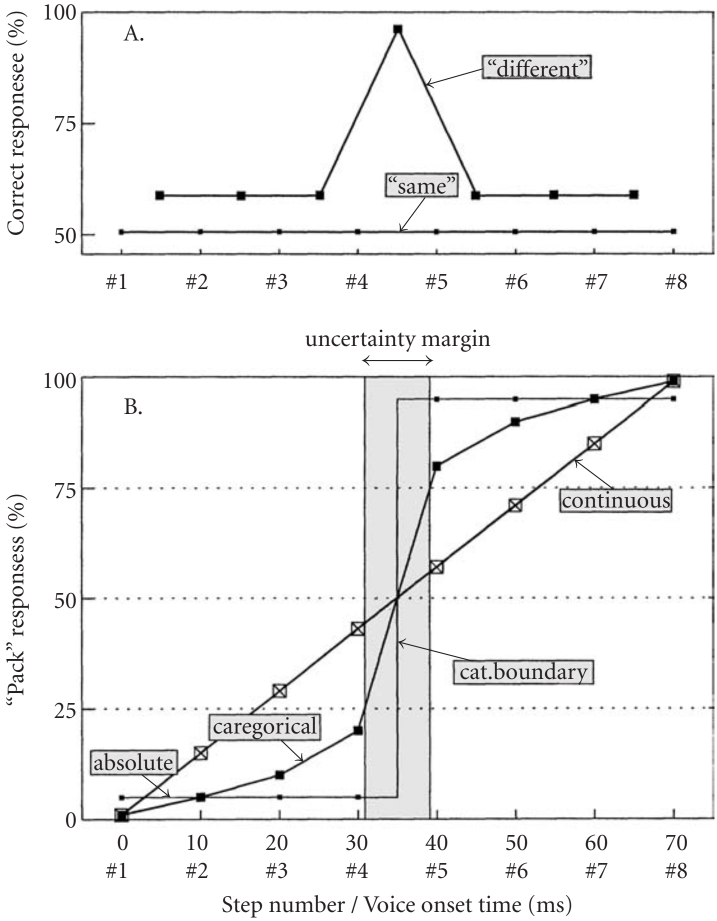


Figure 1. Panel A. Hypothetical discrimination function for physically same and different pairs of stimuli (one-step difference) reflecting categorical perception. Panel B. Illustration of continuous (open squares) versus categorical (big solid squares) perception in the identification and discrimination paradigm. The thin line with small squares represents the ideal step function that should be obtained when categorical perception is absolute. Category boundary and uncertainty margins are indicated (further, see text).

The results of a discrimination task are best expressed as the percentage of correct decisions obtained for a different stimulus pair minus the percentage of errors for same pairs constructed from these stimuli (the latter percentage is often called the response bias). In the case of true categorical perception the discrimination scores show a pronounced peak for the stimulus pair straddling the category boundary, whilst all other pairs are discriminated at or only little above chance level (see panel A in Figure 1). Physically different sounds that fall in the same perceptual category are hard to discriminate. In the case of continuous perception, there is no local peak in the discrimination scores.

1.2 Categorical nature of intonational contrasts

By intonation or speech melody we mean the pattern of rises and falls in the course of the pitch of spoken sentences. Melodic patterns in speech vary systematically across languages, and even within languages across dialects. The cross-linguistic differences can be parametrized and described in much the same way as has been done for the segmentals in language: a set of distinctive features defines an inventory of abstract units, which can be organised in higher-order units subject to wellformedness constraints. Moreover, intonational contrasts are used to perform grammatical functions that can also be expressed by lexico-syntactic means, such as turning statements into questions or into commands, and putting constituents in focus. For these reasons it has become widely accepted that intonation is part of the linguistic system (Ladd 1996: 8). Yet, there have always been adherents of the view that speech melody should be considered as something outside the realm of linguistics proper, i.e., that intonation is a paralinguistic phenomenon at best, to be treated on a par with the expression of attitudes or emotions. Typically, the communication of emotions (such as anger, fear, joy, surprise) or of attitudes (such as sarcasm) is non-categorical: the speaker shows himself more or less angry, fearful, or sarcastic in a continuously gradient fashion.

A relatively recent insight is that a division should be made in melodic phenomena occurring in speech between linguistic versus paralinguistic contrasts. Obviously, only the former but not the latter type of phenomena should be described by the grammar and explained by linguistic theory. This, however, begs the question how the difference can be made between linguistic and paralinguistic phenomena within the realm of speech melody.[2] A methodology to decide on the linguistic status a a melodic phenomenon, i.e., to decide whether the phenomenon should or should not be treated in the grammar, has only recently begun to develop.

Consequently, the objective of the present study is to gain insight into the the psycholinguistic nature of intonational distinctions and to apply this knowledge to the development of formal procedures to establish the linguistic nature of hypothesized categories in intonational systems. We will report an experiment that involves the use of the categorical perception paradigm to a contrast of boundary tones in Dutch. Assuming that these boundary tones constitute a undisputed binary categorical contrast in the language, we can determine whether the categorical perception paradigm is sensitive at all to serve as a diagnostic tool in settling the linguistic status of a melodic parameter.

2. Experimental approach – categorical perception of boundary tones

For segmentals, the aforementioned categorical perception (CP) paradigm (Liberman et al. 1957) can be used to falsify hypothesized phonemic distinctions. This paradigm involves two perceptual tasks: first, listeners classify utterances as exponents of one the two hypothesized categories. After that the listeners are subjected to a discrimination task: they are asked to discriminate stimuli that are either the same or slightly different. Language users are more sensitive to acoustic differences between categories than within categories. Perception is deemed to be categorical when the position of the boundary between the two hypothesized categories in the classification responses coincides with a peak in discriminatory precision (cf. §1.1).[3]

Ladd and Morton (1997) were the first to apply this method to intonation; they investigated a contrast between two pitch-accents in English, viz. the regular high pitch-accent and the emphatic high, which involves a larger excursion size. This distinction can be analysed as a matter of paralinguistic variation within the single phonological category of accent (Ladd & Terken 1995). The alternative hypothesis is that the emphatic high is a linguistic category distinct from the normal high. The classification experiments of emphatic versus normal accents indeed revealed S-shaped curves; however, the subsequent discrimination experiments failed to show the required peak in discriminatory precision at the category boundary.

The ambiguity involved in the interpretation of Ladd and Morton's results is that it is unclear what we are to expect when applying the CP paradigm to intonation: categorical perception in its strongest form is characteristic for stop consonants only; there the peak in discriminatory precision can be predicted from the classification results (Schouten & van Hessen 1992). When the method is applied to vowel contrasts, the improvement of discriminatory precision at the crossover is less clear, e.g. taking the shape of a distributed plateau rather than that of a local peak (Schouten & van Hessen 1992). Acoustically, intonational units are similar to vowel phonemes in that they are encoded over relatively long stretches of speech signal; as such they may be equally perceived in continuous rather than categorical manner. Consequently, when the CP paradigm is applied to an intonational contrast, various correspondence patterns between classification and discrimination results may emerge: a peak or a plateau at the crossover signals categorical perception. In the case of Ladd and Morton (1997), however, neither pattern was encountered, which leaves two possibilities: (i) the normal versus emphatic distinction could be considered as paralinguistic, or (ii) it may be that the CP paradigm is inadequate for distinguishing intonational contrasts. Option (ii) is the position taken by Ladd and Morton (1997) themselves, who maintain that the distinction should be considered as linguistic:

> That is, it does not seem unreasonable to suggest (...) that listeners are predisposed to interpret accents or utterances as being categorically either 'normal' or 'emphatic'. A variety of acoustic and pragmatic parameters play a role in this decision (...). [T]hese parameters may be continuously variable, and the continuous variability may be directly perceptible as such, and there is thus no truly categorical perception. Yet the interpretation computed on the basis of all the input parameters nevertheless normally falls unambiguously into one category or the other. (Ladd & Morton 1997: 339)

In order to decide between the alternative interpretations, we should first find out whether the CP paradigm is applicable to intonation. We can do so by investigating whether the criterion for categorical perception – crossover point corresponding with a peak in discriminatory precision – is satisfied when the CP paradigm is applied to an intonational contrast that is linguistic beyond suspicion. If such a contrast is perceived categorically, then the normal versus emphatic distinction should be interpreted as paralinguistic. If not, Ladd and Morton's (1997) interpretation of this contrast (see citation above) may be relevant for intonational contrasts in general. This plan is carried out in the experiment below, using the contrast between a low (L%) vs. a high (H%) boundary tone in Dutch. H% signals 'appeal to the listener' (Kirsner, van Heuven, & van Bezooijen 1994) and is the prototypical question marker ('t Hart, Collier, & Cohen 1990; Gooskens & van Heuven 1995; van Heuven & Haan 2000); L% marks the end of a statement.

3. Method

3.1 Stimuli

The utterance selected (1) is segmentally ambiguous between a statement and a question, the contrast being encoded by F_0. In both sentences, there is a high pitch-accent on /dɛn/, the lexically stressed first syllable of the noun.[4] While in the statement F_0 remains low (L-L%), the question realisation features a late rise in the final syllable (L-H%):

(1) *De Dennenlaan.*
/də dɛnəlaːn/
'Pine Lane.'

The source sentence was realised by a state-of-the-art Dutch diphone text-to-speech system (Fluency Speech Technology, 1997), with a male voice, and subsequent manipulation of pitch (or rather: its acoustical correlate fundamental frequency, or F_0) in order to vary the F_0 contour from a question realisation to a statement (see Figure 2). Although in natural speech intensity may co-vary with F_0, we kept intensity constant in order to isolate the perceptual effect of F_0, which we consider the primary correlate of intonation. The F_0 contours

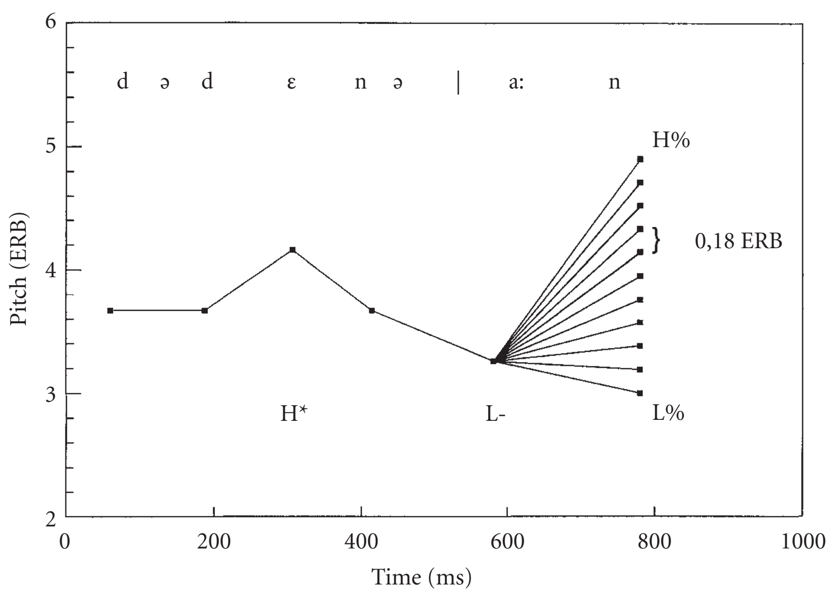


Figure 2. The stimulus F_0 continuum (11 steps) between L-L% and L-H%.

were generated using straight-line interpolation in the PSOLA resynthesis option provided by the Praat speech analysis and resynthesis software (Boersma & Weenink 1996).

Varying the terminal pitch point in Figure 2 triggers the perceptual variation of sentence type. Once the extremes of the continuum had been established,[5] an 11-step continuum was created, at equidistant[6] points along the ERB scale, using the hertz-to-ERB conversion formula provided by the Praat software.[7] The stimuli were numbered 1 (clear statement) to 11 (clear question).

3.2 Tasks and experimental procedures

The tasks and the experimental procedures follow Ladd and Morton's (1997) approach: the classification task is a forced choice setup in which the listeners respond to a stimulus from the continuum by classifying it as either a question or a statement. There were ten repetitions of each of the 11 points from the continuum. These 110 stimuli were presented in random order.

For the discrimination task, we followed Ladd and Morton (1997) in using the AX discrimination paradigm. Stimuli were presented in pairs that are either the same or one step apart on the continuum. In the latter case, the second can be higher or lower than the first (hereafter resp. AB and BA). The ten AB stimulus types ran from pair {1,2} to {10,11}; the ten corresponding BA types from {2,1} to {11,10}. There were five tokens of each pair. This yielded 200 discrimination stimuli (50 AB, 50 BA, 100 AA).

Seventeen native Dutch listeners, 10 males and 7 females, took part in the experiment on a voluntary basis. Participants were university students and staff from various departments at Leiden University. None of them reported any perceptual deficiencies. They were paid *f* 15 for their services. We only took into account the responses of 13 listeners.[8]

The experiments were run individually or in small groups of subjects, who listened to the stimuli over headphones, while seated in a quiet lecture room. Subjects marked their responses on printed answer sheets provided to them, always taking the identification task first and the discrimination task later. For both tests, the actual stimuli were preceded by a short practice session, so that the participants could become acquainted with the full range of stimuli. The whole experiment took about 55 minutes.

4. Results

4.1 Classification task

Figure 3 illustrates the perceptual crossover between the two categories. The psychometric function is unmistakably S-shaped, going from less than 20 to more than 80% ‘statement’ responses within three steps of the 11-point continuum. A probit analysis (Aldrich & Nelson 1984) suggests that the exact crossover point between ‘question’ and ‘statement’ lies at an (interpolated) terminal F_0-value of 122 Hz.

In general, the identification curves of individual subjects are very steep, crossing over (from 25% to 75%) in just one or two steps (see Figure 4). This precision is obscured in Figure 3, where the data are averaged over subjects: clearly, there is substantial between-subject variability in the location of the crossover; a stimulus of which all tokens are marked as questions by one subject may be consistently associated with declarative meaning by another subject. The large between-subject variation is unexpected, differing from the situation encountered for segmentals (Schouten, p.c.).[9] This phenomenon has considerable influence on the data analysis, since it predicts by necessity a high-discrimination plateau rather than a local discrimination peak.

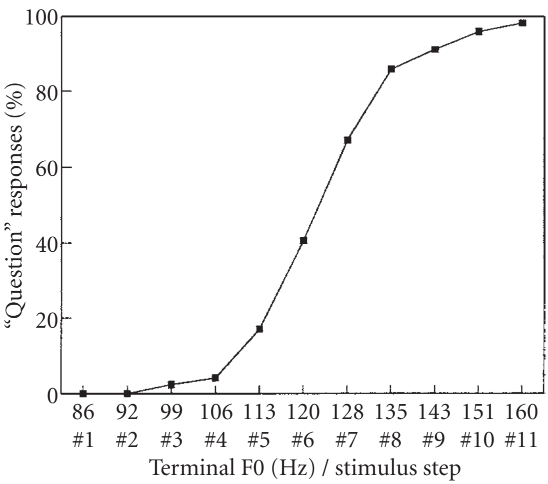


Figure 3. Percent ‘question’ responses as a function of stimulus step number. Data have been accumulated across all listeners.

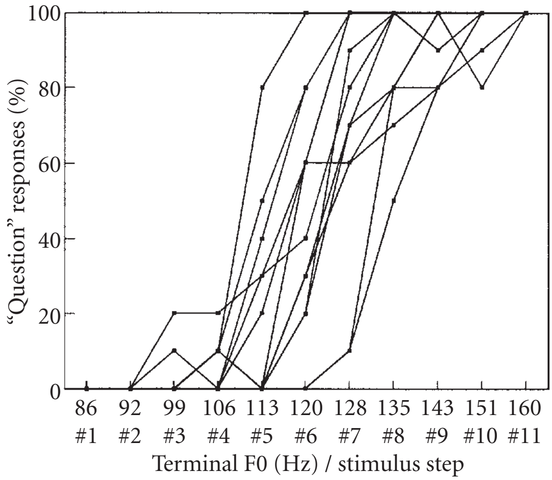


Figure 4. Percent 'question' responses as a function of stimulus step number. Data are presented for individual listeners.

4.2 Discrimination task

Figure 5 presents the mean percentage of succesfully discriminated different stimuli (hereafter 'hits'), and the percentage of false alarms ('different' responses to – identical – AA stimuli). Various aspects of these results will be discussed in the following sections.

4.2.1 *An unexpected peak in discriminatory precision*

Both AB and BA hit curves reach a maximum at the low end of the stimulus range (see Figure 5). Yet, this phenomenon should not be interpreted as an artifact of ERB scaling. Had this been the case, the increase would have been gradual rather than sudden. Therefore, the explanation is to be sought in the domain of linguistics rather than that of psychophysics. Stimuli that are consistently marked as statements are discriminated relatively better. Possibly, the continuum we presented actually reflects three linguistic categories instead of two. The evidence is inconclusive as the continuum ends within the maximum. Moreover, we would need other semantic labels in the classification task in order to test this distinction by means of the CP paradigm. The hypothetical categorical distinction may crucially depend on the amount of final lowering.

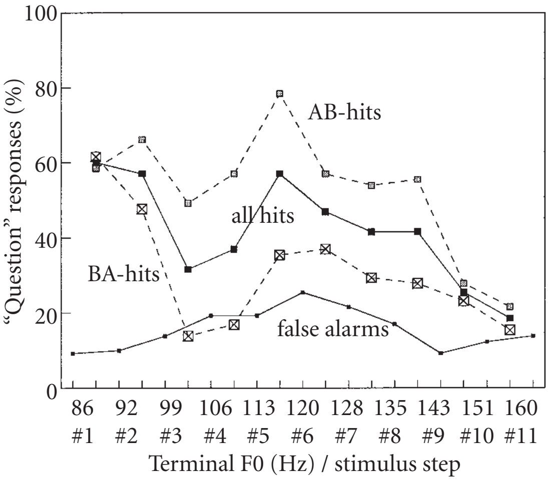


Figure 5. Summary of discrimination results. Percent hits (AB and BA-orders presented separately as well as averaged) and false alarms as a function of stimulus step number (terminal F_0 of boundary tone).

4.2.2 *Order of presentation effect*

There is an order of presentation effect in the discrimination data (see Figure 5): two differrent contours are more succesfully discriminated when the second one has the higher terminal pitch (AB sequence). The same effect has been observed for pitch-accents by Ladd and Morton (1997). They found that the size of the effect depends on step size: when the difference between A and B is minimal, the hit rate for the BA stimuli was equal to the false alarm rate (i.e., chance performance). As step size is increased, more BA stimuli could be distinguished, but never as many as AB stimuli. In our data, the BA hit curve lies between the curve for false alarms and the one for AB hits. We tend agree with Ladd and Morton's interpretation, that the phenomenon is related to declination: '(...) the second accent peak in an utterance is perceived to be equally prominent when it is actually slightly lower in F_0' (Ladd & Morton 1997: 331). Though this effect was originally observed for sequences of pitch-accents within an intonational phrase (Pierrehumbert 1979), Ladd and Morton found it to occur across a phrase boundary; our data suggest it extends to phrase tones.

4.2.3 *The category crossover and the discrimination peak*

Ignoring the peak at the low end, both the AB and BA curves show a second maximum in the middle of the continuum, between 113 and 128 Hz (Figure 5). The BA curve reaches a plateau between 116 and 124 Hz, with the value at the 124 Hz data point slightly higher. The maximum of the AB curve is located at 116 Hz.

There are various ways to check whether these peaks correspond with the crossover between the categories in the classification task. One is to apply the so-called Haskins formula,[10] resulting in a curve of discrimination as predicted by the classification results. This curve can then be compared with the actual discrimination responses, and this comparison forms the basis for deciding whether the contrast can be interpreted as categorically perceived. The predicted discrimination curve features a plateau-like maximum, with a peak at 124 Hz. This corresponds to the BA maximum; also, the plateau extends one step to the left, where the peak of the AB curve is located.

An alternative procedure is used by Ladd and Morton (1997): they compute the point by means of a probit analysis. In our case, this results in a value of 122 Hz, which is in between the AB and BA maxima.

When discussing the classification results, we found a large amount of between-subject variability. From this fact we may derive yet another heuristic for categorical perception of the boundary tone contrast under investigation; if the crossover varies between subjects, then the CP theory predicts the location of individual maxima in discriminatory precision to vary accordingly. Therefore, we performed a regression analysis with the individual crossover points as predicted by probit analysis as the predictor variable and the individual peak in the discrimination results (AB and BA hits collapsed), i.e., the point on the continuum for which most hits were marked, ignoring the responses below 106 Hz, as the criterion.[11] Figure 6 shows the correlation in a scatterplot. The correlation is significant ($r = 0.68$, $p = 0.02$). Therefore, while subjects vary in perceiving the location of the hypothesized category boundary (Figure 3), the responses of the individual subjects feature a peak in discriminatory precision that significantly correlates with the location where they perceive the category boundary.

5. Conclusion

We have accumulated evidence for the phonological status of the contrast between boundary tones in Dutch. In accordance with CP theory (Liberman et

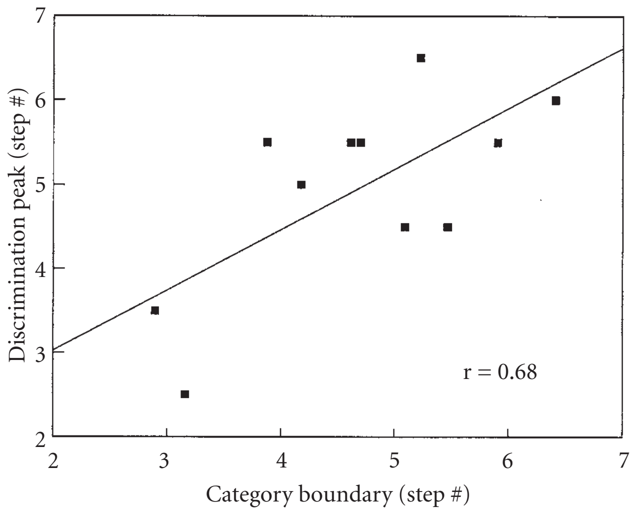


Figure 6. Scatterplot showing the relation between the peak in the individual discrimination function (averaged over AB and BA orders) and the location of the individual cross-over in the identification task (as predicted by probit analysis).

al. 1957), discriminatory precision is better at the category boundary.[12] The large amount of between-subject variation has lead us to diagnose this correspondence in an alternative way, namely by means of a regression analysis. If the large between-subject variation observed is a general characteristic for intonational categorical distinctions, this heuristic may prove fruitful for future research. For example, it may be that Ladd and Morton's (1997) can be interpreted in a different manner by taking into account between-subject variability. If not, we would conclude that their emphatic versus normal contrast is non-linguistic.

Ladd and Morton (1997) concluded that the normal versus emphatic contrast may be categorically interpreted while still being perceived continuously (see §1.2). Our data are in disagreement with this hypothesis: our results provide a clear instance of categorical perception of an intonational contrast. A problem with the Ladd and Morton (1997) study may have been the rather small stepsize of the F_0 increments: their discrimination scores were essentially random. The authors state in a footnote that they ran pilot versions of the experiment with doubled stepsizes, for which they found near-perfect discrimination across the entire continuum, and were therefore content with maintaining the original one-step increments. The stepsize used in our experiment lies halfway the one-step and the two-step increments of Ladd and Morton. We would argue that our stepsize roughly coincides with the difference limen for F_0(-change) in speech, whilst the Ladd and Morton one-step increments were subliminal. Our data, therefore, show that the classical combination of identification and discrimination provides a feasible heuristic after all to decide whether a melodic contrast is linguistic (i.e., categorical, quantal) or paralinguistic (i.e., continuous, scalar).

However, to end this section on a more negative note, our experiment also shows the shortcomings of the classical CP as a tool in intonation research:

– Firstly there is the above-mentioned meaning-dependency of the classification task; even in a relatively clear case such as this boundary tone contrast, it is still possible for subjects to relate the stimuli to more/other categories than the ones offered (see Note 8, see also §6 below).
– Next, phonological contrasts in intonation do not always correspond to a pragmatic contrast by themselves: they acquire part of their meaning syntagmatically (see Note 4). It is therefore problematic to link a section of the contour to the pragmatic content of the whole contour in a classification experiment.
– Finally, the discrimination task is obviously very hard for subjects, since some may find all stimuli to lie below the perceptual threshold (see Note 8). Moreover, subjects professed having had trouble staying concentrated.

We observed for boundary tones the same order of presentation effect as Ladd and Morton (1997) did for pitch-accents. On the basis of these results, we conclude that the BA order of presentation can be neglected in discrimination experiments on intonation, as the effect lowers the discriminatory threshold. Consequently, one can obtain the same results with half of the stimulus set.

Unexpectedly, we encountered a peak in discriminatory precision at the low end of the continuum. This may indicate the presence of another linguistic category which is acoustically characterized by final lowering. Even though the pragmatic meaning of final lowering is much less clear than the question versus declarative contrast associated with a high boundary, it may equally be linguistic and therefore enhance discrimination. Pragmatic interpretations of this category would be speculative – this confirms that it is preferable to diagnose intonational categories meaning-independently.[13]

6. Prospect – a better heuristic?

Even though our CP results can be interpreted as a confirmation of the categorical nature of the boundary tone contrast under investigation, there are reasons to look for alternative procedures. For one thing, the classification task is too dependent on meaning; when CP is applied to a segmental contrast such as /b/ versus /p/, the contrast can operationalised as a difference in lexical meaning, as in the minimal word pair *back* versus *pack*. But using CP in intonational phonology, one has to label the categories by referring to their pragmatic effect. Unfortunately, meaning is highly elusive in the majority of the functions expressed by intonational means. What semantic label, for instance, can one give to the high (H*) versus low accent (L*) contrast? Even for a relatively straightforward melodic contrast such as the one between the L-L% versus H-H% boundary, there was some controversy on the pragmatic categories exemplified by the stimuli (see Note 8). It seems to us that native listeners have better awareness of linguistic categories that they can readily stick a semantic label to. Since this possibility is not available in the case of most intonational contrasts, listeners will generally feel awkward and reluctant when performing CP on intonational contrasts. What is needed, therefore, is a method that is meaning-independent, and yet addresses the categorical, i.e., linguistic, status of intonational contrasts.

One fruitful avenue would be to appeal to a speaker's mimicking abilities. The first to attempt this approach were Chistovich et al. (1966), who had one speaker (Chistovich herself) imitate as well as she could the vowel quality of a series of twelve acoustically equidistant (in terms of F_1, F_2, and F_3) vowels sampled from an artificial /i/ to /a/ continuum, once while shadowing at an extremely short latency (i.e., imitation was produced with a 140 ms stimulus-response latency) and once at a more leasurely 900 ms latency. In the short-latency condition the stimulus-response relationship was essentially 1:1, the vowel quality in the imitation perfectly matched that of the model for each of the twelve vowel types. However, in the long-latency condition the imitations tended to cluster (i.e., showed a many-to-few mapping). Chistovich et al. claim that the clusters coincide with the set of extrinsic allophones available to the speaker. Interestingly, Chistovich et al. (1966) also mention the possibility that (some of) of the clusters of mimicked vowels represent vowel phonemes (rather than allophones) in foreign languages that the speaker had access to. And indeed, Schouten (1975) demonstrated the coincidence of such clusters with English vowel phonemes in the mimicking performance of a group of highly advanced Dutch L2 speakers of British English. Let us therefore assume, for the moment, that the clusters of mimicked vowels represent phonemic, rather than allophonic, categories (possibly drawn from multiple languages known to the speaker). We can extend the argument to the imitation of prosodic continua. If we instruct speakers to mimic (at long latency) the eleven exemplars sampled from our L%–H% continuum, their performance should reveal (a tendency towards) clustering in two groups, coinciding with the prototypical L% and H% categories. However, if the speakers mimic exemplars drawn from a non-linguistic (or para-linguistic), continuum, no (or much less) clustering is expected in the response vowels. Obviously, the mimicking task does not explicitly require that the experimental subject attend to categories. It therefore presents an ideal alternative, if indeed it works, to the categorical perception paradigm.

Part of such an intonational mimicking experiment has been undertaken by Pierrehumbert and Steele (1989). They wanted to establish the phonological status of an intonational contrast signalled formally by a difference in tonal alignment. They varied the alignment of the rise-fall sequence in a rise-fall-rise contour, from an early aligned pitch-accent to one that is late-aligned. Five subjects imitated the stimuli from this continuum. The distribution of alignment patterns in the responses tended towards bimodality, i.e., showed two peaks rather than one. Unfortunately, the tendency was not equally present among all subjects: it was absent in one subject's response data, and varied strongly for the other subjects.

Such variablity in the response patterns across individual speakers was apparent in earlier imitation experiments on segmental continua by Kent (1973, 1974) on vowel glides and by Schouten (1975, 1977) on stationary vowels. These authors concur in their explanation that the imitation process involves (at least) two kinds of memory representation: (i) an immediate pre-categorical sensory storage, and (ii) a categorical storage that contains the output of a phonetic feature analysis performed on the contents of (i). When under extreme time pressure, e.g. when shadowing at short latencies, the subject may choose to skip stage (ii), and feed the output of (i) directly to the motor system.[14] If this explanation is correct, the mimicking technique should be limited to the long-latency condition only (i.e., no time pressure), and even then speakers should be screened individually post hoc.

Human speakers differ widely in their overall pitch. Due to anatomic-physiological reasons, the vocal cord vibration (the primary articulatory determinant of vocal pitch) of female voices is about twice as fast as that of males, and even within the sexes average pitch varies substantially between individuals (even if pathological voices are excluded). Kraayeveld (1997: 167) lists mean F_0 for 25 adult Dutch male voices (read-out isolated sentences) between 98 and 156 Hz, and for a comparable group of 25 females between 185 ad 264 Hz. Similar effects have been reported for other languages (Hollien & Jackson 1973 for American English; Jassem, Steffan-Batóg, & Czajka 1973 for Polish). As a result of interindividual differences in average pitch, some of the five speakers in the study by Pierrehumbert and Steele (1989) had natural pitches that were closer to the average pitch of the model sentence than others. We do not know to what extent imitation of intonation patterns is affected by the difference in overall pitch between model speaker and imitator. Be this as it may, this circumstance may have contributed to the inconsistency of categorical effects in the mimicking performance of the five speakers. When adopting the categorical mimicking (CM) method, special precautions should be taken to rule out the potential artifact of interspeaker differences in average pitch. For instance, it would make sense to select speakers within one sex only; this reduces the variability between speakers to less than 50% (see above). Next, the stimulus presentation system should probably be made speaker-adaptive, such that the basic stimulus is generated so as coincide precisely within the preferred, natural pitch of each individual speaker.

It should be noted, finally, that Pierrehumbert and Steele only examined a (hypothetical) linguistic contrast. In order to demonstrate the feasibility of the mimicking paradigm as a diagnostic tool to decide between linguistic versus paralinguistic status of an intonational contrast, two reference conditions should be incorporated in the experiment:

– one involving a well-established linguistic categorical use of a melodic contrast, e.g., the L% versus H% boundary tone in Dutch, and
– one varying a tonal dimension communicating undisputedly non-linguistic (i.e., para-linguistic or extra-linguistic) information, such as the continuous use of higher mean F_0 to communicate friendliness (e.g., as employed when talking to infants) or even exasperation and/or dispair.[15]

We know from earlier experiments that the results of the reference conditions will not be absolute. We will never find completely quantal behaviour in the imitations of the undisputedly linguistic dimension, nor will we observe completely scalar mimicking of the tokens sampled from the non-linguistic continuum. Yet the results obtained with these reference conditions define a range within which any target contrast can be located. When the target contrast lies more towards the categorical end of the range we will classify it as linguistic, when it lies closer to the continuous end, the parameter will be considered non-linguistic. Such experiments with the CM methodology are currently under way in our laboratory, and will be reported on in the near future.

Notes

* We thank the following people, whose comments and advice have been very helpful: Johanneke Caspers, Judith Haan, Bob Ladd and Bert Schouten. This research was funded by the Royal Netherlands Academy of Arts and Sciences (KNAW) under project # 95-CS-05 (principal investigators W. A. L. Stokhof & V. J. van Heuven).

1. This commutation procedure is best viewed as a thought experiment; when the exchange is implemented through actual digital tape splicing, the result is more often than not an uninterpretable stream of sound.

2. The nature of the distinction between intonational categories is problematic for a further reason: inter-listener agreement on the identity of intonational events is low (Pitrelli et al. 1994), particularly in comparison with the self-evident consensus on segmental distinctions. This lack of consistency has led Taylor (1998) to reject a basic principle of (intonational) phonology, namely its categorical nature. With respect to methodology, researchers tend to act as expert listeners, linking contours that sound distinct to pragmatic meaning in an intuitive fashion. Accordingly, inter-researcher agreement may be low, too (e.g. Caspers 1998).

3. Since there is no standard procedure to quantify this correspondance, there is a source of subjectivity in the interpretation of CP results (van Hessen & Schouten, forthcoming).

4. In the question realisation, speakers may encode a late-aligned accent: in an exploratory study of natural realizations of the source sentence preceding the perception experiment, all tokens of the two sentence types could be classified succesfully in a Linear Discriminant Analysis on the basis of the F_0 peak in the accented syllable. Apparently, H% is not an independent question marker, but is syntagmatically accompanied by a late-aligned accent. However, the early-aligned accent sounds perfectly acceptable both in the statement and in the question reading, and that is the accent type we will use for stimulus production.

5. The range of the last point in the pitch contour is 70 hertz (Hz).

6. Inter-stimulus distance with respect to the last F_0 point varies from 6.6 Hz at the low end to 7.9 Hz at the high end of the continuum.

7. The equivalent rectangular bandwidth (ERB) scale is currently held to be the most adequate perceptual representation of pitch intervals in an intonation language (cf. Hermes & van Gestel 1991; Terken & Ladd 1995).

8. The responses from three listeners were excluded as the perceptual threshold proved too high for them: they obtained less than 10 percent hits in the discrimination task. One more speaker was excluded because he spontaneously reported more and other pragmatic meanings than 'statement' versus 'question'.

9. However, large between-listener variability has been reported in the cuing of the voiced/voiceless contrast by the duration of the pre-burst silent interval: the boundary was at 70 ms for subject #1 and over 100 ms for subject #7 (Slis & Cohen 1969). These results are commented on by Nooteboom and Cohen (1976: 84) as follows: 'Although the cross-over from /d/ to /t/ proceeds rather gradually when averaged over all listeners, the boundary is quite sharply defined for individual listeners' (our translation, BR&VH). We do not rule out the possibility that the phenomenon of large between-listener variability in cross-overs is

typical of melodic and temporal (rather than prosodic) phenomena, whilst it is absent in spectral (rather than segmental) contrasts.

10. $P(c) = 0.5 \times [1 + (p_1 - p_2)^2]$, where P(c) is the probability of correct discrimination, p_1 is the probability of assigning Stimulus A to one of the categories, and p_2 is the probability of assigning Stimulus B to that same category (see Cutting 1982).

11. When two adjacent points have the highest value, the mean is taken as the data point. For two subjects there was no clear maximum, as two non-adjacent points have the highest value; they were excluded. This reduced the dataset to 11 cases.

12. It was remarked during the HILP-4 conference that categorical perception should not be used as a heuristic for the discrete linguistic nature of a contrast on the strength of the argument that categorical perception is never found for vowel contrasts. Although it is true that categorical perception was traditionally found to be weaker with vowel contrasts than with consonants, more recent research (Schouten & van Hessen 1992; van Hessen & Schouten 1999) has convincingly shown that categorical effects (equal in strength to those reported by Liberman et al. 1957) are obtained for vowels and consonants alike, provided that the stimulus materials are of high quality. Our own stimulus material was manipulated through PSOLA resynthesis, which affords extremely good speech quality.

13. For the phonological status of final lowering: see Liberman and Pierrehumbert (1984).

14. Curiously enough, this prediction proved consistent with the results obtained by Chistovich et al. (1966a, b) and by Schouten (1975, 1977), but is clearly contradicted by Kent (1973, 1974), who found continuous mimicking for long-latency imitation and categorical imitation at short-latency shadowing.

15. In hindsight, a similar non-linguistic reference condition could have been fruitfully included in our own categorical perception experiment as well.

References

Aldrich, J. H., & Nelson, F. D. (1984). *Linear probablility, logit, and probit models*. Quantitative Applications in the Social Sciences, Vol. 45. Sage Publications.

Boersma, P., & Weenink, D. (1996). PRAAT: A system for doing phonetics by computer. Report of the Institute of Phonetic Sciences of the University of Amsterdam 132. [http://fonsg3.let.uva.nl/praat]

Caspers, J. (1998). Who's next? The melodic marking of question versus continuation in Dutch. *Language and Speech, 41*, 374–398.

Chistovich, L. A., Fant, G., De Serpa-Leitão, A., & Tjernlund, P. (1966). Mimicking of synthetic vowels. *Speech Transmission Laboratory Quarterly Progress Report on Speech Research*, 1966/2, 1–18.

Chistovich, L. A., Fant, G., & De Serpa-Leitão, A. (1966). Mimicking and perception of synthetic vowels, Part II. *Speech Transmission Laboratory Quarterly Progress Report on Speech Research*, 1966/3, 1–3.

Cutting, J. E. (1982). Plucks and bows are categorically perceived (sometimes). *Perception and Psychophysics, 31*, 462–476.

Fluency Speech Technology (1997). *Fluent Dutch Text-To-Speech. Professional version 1.0*. Utrecht.

Gooskens, Ch., & Heuven, V. J. van (1995). Declination in Dutch and Danish. *Proceedings of the 13th International Congress of Phonetic Sciences*, 374–377. Stockholm.

Grabe, E., Gussenhoven, C., Haan, J., Marsi, E., & Post, B. (1998). Preaccentual pitch and speaker attitude in Dutch. *Language and Speech, 41*, 63–85.

Hart, J. 't, Collier, R., & Cohen, A. (1990). *A perceptual study of intonation. An experimental-phonetic approach to speech perception*. Cambridge: Cambridge University Press.

Hermes, D. J., & Gestel, J. C. van (1991). The frequency scale of speech intonation. *Journal of the Acoustical Society of America, 90*, 97–102.

Hessen, A. J. van, & Schouten, M. E. H. (1999). Categorical perception as a function of stimulus quality. *Phonetica, 56*, 56–72.

Heuven, V. J. van (1986). Some acoustic characteristics and perceptual consequences of foreign accent in Dutch spoken by Turkish immigrant workers. In J. van Oosten, J. F. Snapper (Eds.), *Dutch Linguistics at Berkeley, papers presented at the Dutch Linguistics Colloquium held at the University of California, Berkeley on November 9th, 1985* (pp. 67–84). Berkeley: The Dutch Studies Program, U.C. Berkeley.

Heuven, V. J. van, & Haan, J. (2000). Phonetic correlates of statement versus question intonation in Dutch. In A. Botinis (Ed.), *Intonation: Analysis, modelling and technology* (pp. 119–144). Dordrecht: Kluwer.

Hollien, H., & Jackson, B. (1973). Normative data on the speaking fundamental frequency characteristics of young adult males, *Journal of Phonetics, 1*, 117–120.

Jassem, W., Steffan-Batóg, M. S., & Czajka, S. (1973). Statistical characteristics of short-term average F0 distributions as personal voice features. In W. Jassem (Ed.), *Speech analysis and synthesis*, 3 (pp. 209–225). Warsaw: Panstwowe Wydawn.

Kent, R. D. (1973). The imitation of synthetic vowels and some implications for speech memory. *Phonetica, 28*, 1–25.

Kraayeveld, H. (1997). Idiosyncracy in prosody. Speaker and speaker group identification in Dutch using melodic and temporal information. Doctoral dissertation, University of Nijmegen.

Kirsner, R. S., Heuven, V. J. van, & Bezooijen, R. van (1994). Interaction of particle and prosody in the interpretation of factual Dutch sentences. In R. Bok-Bennema, C. Cremers (Eds.), *Linguistics in the Netherlands 1994* (pp. 107–118). Amsterdam: John Benjamins.

Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.

Ladd, D. R., & Morton, R. (1997). The perception of intonational emphasis: continuous or categorical? *Journal of Phonetics, 25*, 313–342.

Ladd, D. R., & Terken, J. (1995). Modelling intra- and inter-speaker pitch range variation. *Proceedings of the 13th International Congress of Phonetic Sciences, 2*, 386–389. Stockholm.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 61*, 379–388.

Liberman, M., & Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. In M. Aronoff & R. Oehrle (Eds.), *Language and sound structure* (pp. 157–233). Cambridge, MA: MIT Press.

Nooteboom, S. G., & Cohen, A. (1976). *Spreken en verstaan. Een inleiding tot de experimentele fonetiek* [Speaking and understanding. An introduction to experimental phonetics]. Assen: van Gorcum.

Pierrehumbert, J. (1979). The phonology and phonetics of English intonation. PhD dissertation, Massachusetts Institute of Technology.

Pitrelli, J. F., Beckman, M. E., & Hirschberg, J. (1994). Evaluation of prosodic transcription reliability in the ToBI framework. *Proceedings of the 3rd International Conference on Spoken Language Processing, 1*, 123–126. Yokohama.

Schouten, M. E. H. (1975). Native-language interference in the perception of second-language vowels. Doctoral dissertation, Utrecht University.

Schouten, M. E. H. (1977). Imitation of synthetic vowels by bilinguals. *Journal of Phonetics, 5*, 273–283.

Schouten, M. E. H., & Hessen, A. J. van (1992). Modelling phoneme perception. I: Categorical perception. *Journal of the Acoustical Society of America, 92*, 1841–1855.

Slis, I. H., & Cohen, A. (1969). On the complex regulating the voiced-voiceless distinction, *Language and Speech*, 80–102, 137–155.

Taylor, P. (1998). Analysis and synthesis of intonation using the TILT model. Unpublished manuscript, Centre for Speech Technology Research, University of Edinburgh.

Author index

A

Adelaar, K. A. & J. Prentice 159
Alber, B. 123
Aldrich, J. H. & F. D. Nelson 234
Alieva, N. F., V. D. Arakin, A. K. Ogloblin, & Y. H. Sirk 151f.
Allen, J.
see O'Shaughnessy & Allen
Allen, W. S. 123
Anderson, S. R. 52, 84
Arakin, V. D.
see Alieva, Arakin, Ogloblin, & Sirk
Arvaniti, A., D. R. Ladd, & I. Mennen 202
Avery, J. P. 5f., 27

B

Baart, J. L. G. 153, 179, 186f.
Ball, M. J. 83
Ball, M. J.
see also Code & Ball
Basri, H.
see Mithun & Basri
Bastiaanse, R., D. G. Gilbers, & K. van der Linde 90
Battistella, E. L. 82
Beckman, J. 117
Beckman, M. E.
see Pitrelli, Beckman, & Hirschberg
Béland, R.
see Nespoulous, Joanette, Béland, Caplan, & Lecours; Paradis & Béland
Berg, R. van den, C. Gussenhoven, & T. Rietveld 202
Berinstein, A. E. 153
Bernhardt, B. H. & J. P. Stemberger 48, 49
Bezooijen, R. van
see Kirsner, van Heuven, & van Bezooijen
Blake, B. J. & R. M. W. Dixon 6
Blevins, B. 82, 85
Blevins, J. 6, 116, 123
Bloch, B. 37
Blumstein, S. E. 89f.
Blumstein, S. E., W. E. Cooper, H. Goodglass, S. Statlender, & J. Gottlieb 89f.
Boersma, P. 84
Boersma, P. & D. Weenink 233
Bolinger, D. 153
Boogaart, T.
see Damhuis, Boogaart, in 't Veld, Versteijlen, Schelvis, Bos, & Boves
Booij, G. E. 188
Borowsky, T. 3, 5
Borowsky, T. & M. Harvey 124
Bos, L.
see Damhuis, Boogaart, in 't Veld, Versteijlen, Schelvis, Bos, & Boves
Bosch, L. F. M.
see Streefkerk, Pols, & ten Bosch
Bouwhuis, D. G.
see Swerts, Bouwhuis, & Collier
Boves, L.
see Damhuis, Boogaart, in 't Veld, Versteijlen, Schelvis, Bos, & Boves
Brannen, K.
see Goad & Brannen

Broecke, M. P. R. van den
see van Heuven & van den Broecke
Budwig, N. 84
Bybee, J. 84

C
Cairns, C. E. 43
Cairns, C. E. & M. H. Feinstein 43f., 86, 118
Calabrese, A.
see Romani & Calabrese
Caplan, D. 89
Caplan, D.
see also Nespoulous, Joanette, Béland, Caplan, & Lecours; Nespoulous, Joanette, Ska, Caplan, & Lecours
Carlson, L. 123
Caspers, J. 201f., 205, 208, 243
Caspers, J., V. J. van Heuven, & N. van Zwol 206, 208
Charette, M. 3
Chistovich, L. A. 240
Chistovich, L. A., G. Fant, & A. De Serpa-Leitão 240, 244
Chistovich, L. A., G. Fant, A. De Serpa-Leitão, & P. Tjernlund 240, 244
Chitoran, I. 51
Cho, Y. 12
Chomsky, N. 4, 81
Chomsky, N. & M. Halle 49, 56, 82, 85
Christman, S. S. 102
Clements, G. N. 102
Clements, G. N. & E. V. Hume 5
Clements, G. N. & S. J. Keyser 3
Code, C. 89, 103
Code, C. & M. J. Ball 103
Cohen, A.
see 't Hart, Collier, & Cohen; Nooteboom & Cohen; Slis & Cohen
Cohn, A. 152, 171
Cohn, A. & J. J. McCarthy 152
Cole, J. & C. Kisseberth 149
Cole, J. S.
see Sevald, Dell, & Cole
Collier, R. 203
Collier, R.
see also 't Hart & Collier; 't Hart, Collier, & Cohen; Swerts, Bouwhuis, & Collier
Crosson, B., T. A. Novack, & M. R. Trenerry 90
Cooper, W. E.
see Blumstein, Cooper, Goodglass, Statlender, & Gottlieb
Cutting, J. E. 244
Czajka, S.
see Jassem, Steffan-Batóg, & Czajka

D
Damhuis M., T. Boogaart, C. in 't Veld, M. Versteijlen, W. Schelvis, L. Bos, & L. Boves 180
Dardjowidjojo, S.
see Moeliono & Dardjowidjojo
Dell, G. S.
see Sevald, Dell, & Cole
Demarco, S.
see Hough, Demarco, & Farler
Démonet, J. F. 90
Demuth, K. & E. J. Fee 3, 13f.
De Serpa-Leitão, A.
see Chistovich, Fant, & De Serpa-Leitão; Chistovich, Fant, De Serpa-Leitão, & Tjernlund
Dirksen, A. & H. Quené 179
Dixon, R. M. W.
see Blake & Dixon
Dogil, G. & J. Mayer 103
Donegan, P. J. & D. Stampe 102
Donselaar, M. van
see Kuijpers & van Donselaar
Dresher, B. E. & A. Lahiri 123
Dyck, C. 19, 26

E
Ebing, E. F. 156
Eefting, W. Z. F. 156

Eisner, J. 149
Elenbaas, N. 123
Ellis, A. W. & A. W. Young 90
Elson, B. 12
Everett, D. 139, 140ff., 149
Everett, D. & L. Seki 10

F

Fant, G.
see Chistovich, Fant, & De Serpa-Leitão; Chistovich, Fant, De Serpa-Leitão, & Tjernlund
Farler, D.
see Hough, Demarco, & Farler
Favreau, Y., J.-L. Nespoulous, & A. R. Lecours 89f.
Fee, E. J.
see Demuth & Fee
Feinstein, M. H.
see Cairns & Feinstein
Fey, M. & Gandour, J. 14, 15, 20, 26
Fikkert, P. 3, 13f., 87
Fluency Speech Technology 232
Fry, D. B. 153
Fujimura, O. 31, 39, 41ff., 48, 50, 54ff.

G

Gandour, J.
see Fey & Gandour
Gerrits, E. & B. Schouten 204
Gestel, J. C. van
see Hermes & van Gestel
Giegerich, H. 3
Gilbers, D. G. & D. B. den Ouden 87
Gilbers, D. G.
see also Bastiaanse, Gilbers, & van der Linde
Gnanadesikan, A. 4
Goad, H. 13
Goad, H. & K. Brannen 3, 5, 24ff.
Goedemans, R. W. N., E. van Zanten, & Pacilly, J. 151, 173
Goedemans, R. W. N. & E. van Zanten 173
Goldsmith, J. 9, 84
Goodglass, H.
see Blumstein, Cooper, Goodglass, Statlender, & Gottlieb
Gooskens, Ch. & V. J. van Heuven 231
Gordon, M. 116, 123
Gottlieb, J.
see Blumstein, Cooper, Goodglass, Statlender, & Gottlieb
Grabe, E., C. Gussenhoven, J. Haan, E. Marsi, & B. Post 204, 206, 210
Green, A. 143
Green, T. 149
Green, T. & M. Kenstowicz 128, 137, 143, 149
Griffith, B. C.
see Liberman, Harris, Hoffman, & Griffith
Grijns, C. D. 159
Gussenhoven, C. 201ff., 217
Gussenhoven, C.
see also van den Berg, Gussenhoven, & Rietveld; Grabe, Gussenhoven, Haan, Marsi, & Post; Rietveld & Gussenhoven

H

Haan, J.
see Grabe, Gussenhoven, Haan, Marsi, & Post; van Heuven & Haan
Hagstrom, P. 147
Halim, A. 152, 172
Halle, M. & J.-R. Vergnaud 32, 40, 55, 194
Halle, M.
see also Chomsky & Halle
Hammond, M. 147
Hanson, K. & P. Kiparsky 123
Haraguchi, S. 31f., 40ff.
Harrikari, H. 111, 115, 123
Harris, J. 3, 7, 11f., 85f., 88f., 93, 101
Harris, K. S.
see Liberman, Harris, Hoffman, & Griffith

Hart, J. 't & R. Collier 155, 201
Hart, J. 't, R. Collier, & A. Cohen 201, 231
Harvey, M.
see Borowsky & Harvey
Haugen, E. 55
Hayes, B. P. 111, 113ff., 118, 123, 137, 139f., 148, 171, 194
Helsloot, C. J. 183, 194, 198
Helsloot, C. J. & B. M. Streefkerk 179
Hermes, D. J. & J. C. van Gestel 243
Hessen, A. J. van & M. E. H. Schouten 243f.
Hessen, A. J. van
see also Schouten & van Hessen
Heuven, V. J. van 154f., 165, 226
Heuven, V. J. van & E. van Zanten 171
Heuven, V. J. van & M. P. R. van den Broecke 160
Heuven, V. J. van & J. Haan 231
Heuven, V. J. van
see also Caspers, van Heuven, & van Zwol; Gooskens & van Heuven; Kirsner, van Heuven, & van Bezooijen; Remijsen & van Heuven; van Zanten & van Heuven
Hewitt, M. 127
Hirschberg, J.
see Pierrehumbert & Hirschberg; Pitrelli, Beckman, & Hirschberg
Hoard, J. E. 12, 19, 24
Hoffman, H. S.
see Liberman, Harris, Hoffman, & Griffith
Hollander, J. J. de 152
Hollien, H. & B. Jackson 241
Holman, E. 111
Holmes, U. T. 13f., 18, 26
Honma, T. 48
Hough, M. S., S. Demarco, & D. Farler 89f.
Hulst, H. G. van der 86
Hume, E. V.
see Clements & Hume
Hyman, L. M. 118, 123

I

Ikranagara, K. 159
Ingram, D. 3
Itô, J. 3, 7f., 10, 33
Itô, J. & R. A. Mester 127f., 147

J

Jackson, B.
see Hollien & Jackson
Jakobson, R. 3f., 21, 82f.
Jalava, A. 111
Jassem, W., M. S. Steffan-Batóg, & S. Czajka 241
Jensen, J. 8, 10, 12, 26
Joanette, Y.
see Nespoulous, Joanette, Béland, D. Caplan, & A. R. Lecours; Nespoulous, Joanette, Ska, Caplan, & Lecours

K

Kager, R. 114, 123, 128f., 137, 143
Kager, R.
see also Quené & Kager
Kangasmaa-Minn, E. 111
Karlsson, F. 115
Katayama, M. 56
Kawasaki, T. 12, 27
Kaye, J. D. 3, 6, 64f., 68, 70
Kaye, J. D., J. Lowenstamm, & J.-R. Vergnaud 6f., 60, 63f., 88
Keating, P. A.
see Westbury & Keating
Keijsper, C. E. 205
Kenstowicz, M. 113, 123, 146, 149
Kenstowicz, M.
see also Green & Kenstowicz
Kent, R. D. 241, 244
Key, H. 3, 133, 135ff., 148f.
Keyser, S. J.
see Clements & Keyser
Kiparsky, P. 71

Kiparsky, P.
see also Hanson & Kiparsky
Kirsner, R. S., V. J. van Heuven, & R. van Bezooijen 231
Kisseberth, C.
see Cole & Kisseberth
Kjellin, O. 51
Kohn, S. E. 90
Koreman, J.
see Lahiri & Koreman
Kraayeveld, H. 241
Kruyt, J. 187
Kuijpers, C. & M. van Donselaar 183

L
Ladd, D. R. 153, 201, 203f., 229
Ladd, D. R.
see also Arvaniti, Ladd, & Mennen
Ladd, D. R. & R. Morton 204, 230f., 233, 236ff.
Ladd, D. R. & J. Terken 230, 243
Ladefoged, P. 16
Lahiri, A. & J. Koreman 116
Lahiri, A.
see also Dresher & Lahiri
Laksman, M. 151
Lapoliwa, H. 152
Laver, J. 25, 116
Lea, W. 187
Lecours, A. R.
see Favreau, Nespoulous, & Lecours; Nespoulous, Joanette, Béland, Caplan, & Lecours; Nespoulous, Joanette, Ska, Caplan, & Lecours
Leopold, W. F. 14, 19, 26
Levin, J. 133, 135
Liberman, A. M. 83
Liberman, A. M., K. S. Harris, H. S. Hoffman, & B. C. Griffith 230, 237, 244
Liberman, M. 118
Liberman, M. & A. S. Prince 118, 183
Liberman, M. & J. B. Pierrehumbert 244
Linde, K. van der
see Bastiaanse, Gilbers, & van der Linde
Lombardi, L. 12
Lowenstamm, J. 59, 64, 66, 71, 76, 78
Lowenstamm, J.
see also Kaye, Lowenstamm, & Vergnaud

M
Macken, M. A. 82
Marsi, E.
see Grabe, Gussenhoven, Haan, Marsi, & Post
Mayer, J.
see Dogil & Mayer
McCarthy, J. J. 116, 118
McCarthy, J. J. & A. S. Prince 3, 112, 114, 129, 136
McCarthy, J. J.
see also Cohn & McCarthy
McCartney, S. 113, 123
Meijer, P. J. A. 85
Menn, L. 26
Mennen, I.
see Arvaniti, Ladd, & Mennen
Mester, R. A.
see Itô & Mester
Mithun, M. & H. Basri 9
Moeliono, A. M. & S. Dardjowidjojo 151
Morton, R.
see Ladd & Morton
Moulton, W. G. 55
Muhadjir 159

N
Nababan, P. W. J. 155
Nádasdy, Á. & P. Siptár 59, 67, 68
Nahkola, K. 111
Nelson, F. D.
see Aldrich & Nelson
Nespoulous, J.-L., Y. Joanette, R. Béland, D. Caplan, & A. R. Lecours 89f.

Nespoulous, J.-L., Y. Joanette, B. Ska, D. Caplan, & A. R. Lecours 89f.
Nespoulous, J.-L.
see also Favreau, Nespoulous, & Lecours; Wilshire & Nespoulous
Nespor, M. & I. Vogel 136, 182
Nooteboom, S. G. 156
Nooteboom, S. G. & A. Cohen 243
Novack, T. A.
see Crosson, Novack, & Trenerry

O
Odé, C. 152
Ogloblin, A. K.
see Alieva, Arakin, Ogloblin, & Sirk
O'Shaughnessy, D. 186
O'Shaughnessy, D. & J. Allen 187
Ouden, D. B. & R. Bastiaanse 81
Ouden, D. B. den
see Gilbers & den Ouden

P
Pacilly, J. J. A. 218
Pacilly, J. J. A.
see also Goedemans, van Zanten, & Pacilly
Palander, M. 111
Papp, F. 60
Paradis, C. & R. Béland 87
Paunonen, H. 111
Pierrehumbert, J. B. 236
Pierrehumbert, J. B. & J. Hirschberg 204f.
Pierrehumbert, J. B. & S. A. Steele 203, 217, 241f.
Pierrehumbert, J. B.
see also Liberman & Pierrehumbert
Piggott, G. L. 3, 5ff., 9ff., 24
Pinker, S. 5
Pitrelli, J. F., M. E. Beckman, & J. Hirschberg 243
Poedjosoedarmo, S. 155
Polgárdi, K. 59, 73
Polgárdi, K.
see also Rebrus & Polgárdi
Pols, L. C. W.
see Streefkerk & Pols; Streefkerk, Pols, & ten Bosch
Post, B.
see Grabe, Gussenhoven, Haan, Marsi, & Post
Prentice, J. 152
Prentice, J.
see also Adelaar & Prentice
Prince, A. S. 114ff., 140, 145, 194
Prince, A. S. & P. Smolensky (P&S) 13, 88, 111, 129, 134, 144
Prince, A. S.
see also Liberman & Prince; McCarthy & Prince

Q
Quené, H. & R. Kager 179
Quené, H.
see also Dirksen & Quené

R
Rapola, M. 111f.
Ras, J. J. 155
Rebrus, P. 67f., 77f.
Rebrus, P. & K. Polgárdi 68, 70, 76
Remijsen, B. & V. J. van Heuven 204, 225
Rice, K. 3, 5ff., 9, 22
Rice, K. & P. Avery 5
Rietveld, T. & C. Gussenhoven 203, 204, 210
Rietveld, T.
see also van den Berg, Gussenhoven, & Rietveld
Rifkin, J. I. 127ff.
Ritter, N. A. 78
Romani, C. & A. Calabrese 85, 102
Rosenthall, S. 124

S
Samsuri 152ff.
Sapir, J. D. 9, 12

Schelvis, W.
see Damhuis, Boogaart, in 't Veld, Versteijlen, Schelvis, Bos, & Boves
Schouten, B.
see Gerrits & Schouten
Schouten, M. E. H. 234, 240f., 244
Schouten, M. E. H. & A. J. van Hessen 230f., 244
Schouten, M. E. H.
see also van Hessen & Schouten
Seki, L.
see Everett & Seki
Selkirk, E. O. 183
Sevald, C. A., G. S. Dell, & J. S. Cole 85
Siptár, P.
see Nádasdy & Siptár; Törkenczy & Siptár
Sirk, Y. H.
see Alieva, Arakin, Ogloblin, & Sirk
Sherer, T. 123f.
Sherman, A. 56
Shvachkin, N. K. 21
Ska, B.
see Nespoulous, Joanette, Ska, Caplan, & Lecours
Skousen, R. 123
Slis, I. H. & A. Cohen 243
Smith, B. L.
Smith, N. V. 21, 87
Smolensky, P.
see Prince & Smolensky
Spencer, A. 87
Stafford, R. L. 10
Stampe, D. 4
Stampe, D.
see also Donegan & Stampe
Statlender, S.
see Blumstein, Cooper, Goodglass, Statlender, & Gottlieb
Steele, S. A.
see Pierrehumbert & Steele
Steffan-Batóg, M. S.
see Jassem, Steffan-Batóg, & Czajka
Steinhauer, H. 155, 159
Stemberger, J. P. 3, 85
Stemberger, J. P.
see also Bernhardt & Stemberger
Steriade, D. 3, 12, 14, 25, 123
Streefkerk, B. M. 180
Streefkerk, B. M. & L. C. W. Pols 182
Streefkerk, B. M., L. C. W. Pols, & L. F. M. ten Bosch 183
Swerts, M., D. G. Bouwhuis, & R. Collier 210

T
Taylor, P. 243
Teeuw, A. 151
Terken, J.
see Ladd & Terken
Tjernlund, P.
see Chistovich, Fant, De Serpa-Leitão, & Tjernlund
Thorburn, R. 11, 26
Törkenczy, M. 59, 61, 73
Törkenczy, M. & P. Siptár 63, 74
Trenerry, M. R.
see Crosson, Novack, & Trenerry
Trigo, R. L. 9

U
Urbanczyk, S. 19

V
Veld, C. in 't
see Damhuis, Boogaart, in 't Veld, Versteijlen, Schelvis, Bos, & Boves
Vergnaud, J.-R.
see Halle & Vergnaud; Kaye, Lowenstamm & Vergnaud
Versteijlen, M.
see Damhuis, Boogaart, in 't Veld, Versteijlen, Schelvis, Bos, & Boves
Visch, E. A. M. 188

W
Walker, R. 145
Wallace, S. 159, 172
Weenink, D.
see Boersma & Weenink
Westbury, J. & P. A. Keating 12, 14
Wilshire, C. E. & J.-L. Nespoulous 85
Woodrow, H. 160

Y
Yip, M. 41, 50
Yoshida, S. 8, 75
Young, A. W.
see Ellis & Young

Z
Zanten, E. van 152
Zanten E. van & V. J. van Heuven 151f., 156, 170
Zanten, E. van
see also Goedemans & van Zanten; Goedemans, van Zanten, & Pacilly; van Heuven & van Zanten
Zec, D. 123
Zonneveld, R. van 86
Zubkova, L. G. 152
Zwol, N. van
see Caspers, van Heuven, & van Zwol

Language index

A

Arabic
- Moroccan 64f.
- Palestinian 75, 128, 146

Australian languages 6

Austronesian languages 9f.

B

Barasana 149

Bella Coola 19

C

Capanahua 129

Cayuga 26

Cayuvava 3, 10, 127f., 133ff., 148

Chugach 127

D

Diola-Fogny 7ff., 23f.

Dutch 87, 91, 116, 153, 156, 164, 179, 180, 183, 186ff., 192, 195, 197, 201ff., 210, 217, 225, 230, 232f., 240ff.

English 4, 13, 16, 21, 25ff., 32f., 39ff., 44, 46, 48ff., 60f., 135, 153, 156, 187, 226, 240f.
- British English 240

F

Finnish 111, 113, 116, 122f.
- Eastern and Northern dialects 111

French 8, 11, 27, 129

G

Georgian 51

German 40, 55

H

Hebrew 56

Hungarian 59ff., 63ff., 69ff., 73f., 76, 78, 147

I

Icelandic 52

Indonesian 151ff., 159, 161ff., 165, 169ff.
- Jakartan Indonesian 159, 162, 166, 167ff., 171ff.

Italian 194

J

Japanese 8, 11, 32ff., 37, 39, 49, 54ff., 128
- Tokyo Japanese 33f., 36

Javanese 155ff.

K

Kamaiurá 10f.

L

Latvian 129

Luo 10

M

Malay 159
- Jakarta Malay 159
- *see also* Indonesian

Maranungku 129

N
Nez Perce 12
Norwegian 71

O
Op (ob)-language 48f.

P
Pig Latin 48f.
Pirahã 128, 139ff.
Polish 74, 241
Popoluca, Sierra 12

S
Selayarese 8ff.
Senufo 3, 10
Swedish 51f.

T
Toba Batak 155ff., 162, 165ff.
Turkish 64

W
Weri 129
West Germanic languages 13

Y
Yapese 8ff., 23ff.
Yawelmani 64

Subject index

A
accent 153
  phrase-based accent 170
  phrasal accent 170
  see also pitch accent
accent-lending fall 203ff.
  timing of 203ff.
acquisition 4ff.
allomorphy 38, 41, 56
allophone 225
aphasia 81, 83, 89ff., 101
apraxia of speech 89f., 103
aspiration 4, 8ff.
assimilation 35, 37, 47
AX discrimination 233

B
behaviourism 84
bogus clusters 62f., 76

C
categorical perception 226, 229
category boundary 227
child language 3ff.
closed syllables 60ff.
cluster reduction 87, 101
coda constraint 13, 16, 19, 23, 24, 33
cola 143
complex onsets 61f.
compounds 188ff.
consonant clusters 41, 55
constraints 33
  constraint hierarchy 197
  see also prosodic constraints
content words 180ff.
contrast *see* Theory of Maximal Contrast
contrastive focus 186
Convertor/Distributor Model 31ff.
correlates of stress *see* stress
cue trading 226

D
deletion 89, 93ff., 101
discrimination function 227
  see also AX discrimination; discrimination function
distinctive features (as unordered set) 44

E
epenthesis 8ff., 47, 70, 73f., 76
ERB scale 233
evaluation test 159
extrametricality 144, 146, 148

F
faithfulness, positional 112ff.
fall *see* accent-lending fall
false alarms 235
final length (of consonants) 15ff.
final release 4, 18ff.
'finality' attitude 210
finality vs. non-finality (in sentence) 189ff.
focus 191
foot binarity 13f.
frequency 191
function words 180ff.
functional nativism 84, 103

G
geminates 34f., 46
gemination 111ff.
Government Phonology 6ff., 26, 59ff., 88, 93, 101f.

H
Haskins formula 237
hits 235

I
iambs 111, 190, 195
identity requirement (on proper government) 75
information *see* 'new' information; projected information
intonation 229
intonation models 201ff.

L
language acquisition 4
lapse avoidance 196f.
LAPSE constraint 137, 143ff.
Laryngeal languages 5, 21ff.
laryngeal neutralization 12, 14
lengthening *see* vowel lengthening
lexicalisation 191
lexicon 38
licensing 3, 5, 9ff.
linearization 31ff.
long vowels 59ff.
lowering 70, 74f., 78

M
markedness 82ff., 88, 91, 101, 103
  syllable markedness 112
maximality (of prosodic domain) 193f.
mean F_0 241
meaning difference 204
metrical-prosodic structure 179ff.
metrical grid 183, 187, 194, 198
metrical theory 127
mimicking 240f.
minimal word 115, 128, 148
  see also word minimality
monosyllables vs. polysyllables 184ff.
moraic nasal 34f.
morphological structure 37, 49ff.

N
nasal release 4, 15, 20
nativism 84
  see also functional nativism
'new' information 206

O
Obligatory Contour Principle 38, 47
onset constraint 33
onset height (in intonation) 205
onsets (of empty headed syllables) 3f., 7ff.
  see also complex onsets
Optimality Theory 88, 111

P
pairwise comparison experiments 158
paraphasia *see* phonemic paraphasia
perception experiments 153, 180f.
perception *see* categorical perception
phoneme 225
phonemic paraphasia 89f., 93
phonology (vs. phonetics) 31ff.
phonotactic restrictions 59ff.
phrasal stress 128, 135, 148
pitch
  male vs. female 241
  overall pitch 241
  pitch movement 155
pitch accent 153
place specification 33ff.
pointed hat intonation contour 204
positional faithfulness *see* faithfulness
projected information 207
prominence 168, 179ff.
  perceived prominence 179ff.
prosodic constituency 182ff.
prosodic constraints 183ff.

prosodic hierarchy 127
prosodic parsing 179ff.
prosodic properties 153, 179ff.
prosodic word 127ff.
PSOLA resynthesis 233
psychometric function 227

Q
quantity-sensitivity 111ff.
quaternary stress 147f.

R
reduction *see* cluster reduction
release 4, 6ff.
repetition task 91, 92
response bias 229
resynthesis *see* PSOLA resynthesis
rhythmic alternation 193ff.
rhythmic lapses and clashes 184f., 195f.

S
/s/ plus stop clusters 39
secret (disguised) languages 48
Set Theory (of the syllable) 31ff.
shortening 59, 63f., 68ff., 76, 78
single stress phrase 141f.
sonority 41, 53f., 87ff., 99, 101, 116
Sonority Dispersion Principle 102
speech rate 185
spontaneous voicing 5ff.
spontaneous voicing languages 6ff., 20ff., 27
stress 151ff.
    stress typology 144
    correlates of stress 153
    free stress
strict CV 59ff.
Strict Layering 131, 147
substrate languages 152
superheavy rhymes 61
syllable structure 31ff., 60ff., 81, 85ff., 91ff., 102, 103, 104, 116ff.
syllable
    deletion 183f.
    insertion 183f.
syntactic structure 179

T
templates 86, 88, 93, 101
ternary stress 128, 135, 145, 148
text-to-speech systems 179, 184f., 188, 197, 198
Theory of Maximal Contrast 102
trochees 111ff., 193, 195
    moraic trochees 111
    syllabic trochees 111ff.
truncation 13ff.
typology 4, 8ff.

U
uncertainty margin 227
underparsing 71f., 75, 76, 78
underspecification 31

V
Velar Softening 47
voice assimilation 42, 50f.
voicelessness 50
vowel length 13, 15, 17ff., 112ff.
    see also long vowels
vowel lengthening 59, 67, 70
vowel quality 226

W
weak local parsing 137, 143f.
word minimality 11, 13ff.

Table of contents, volume I

Preface vii

Nasality, voice and more ix

Nasality

Nasal harmony in functional phonology 3
Paul Boersma

Reinterpreting transparency in nasal harmony 37
Rachel Walker

Can 'phonological' nasality be derived from phonetic nasality? 73
Stefan Ploch

Voice

The role of phonology and phonetics in Dutch voice assimilation 119
Mirjam Ernestus

Final Devoicing and the stratification of the lexicon in German 145
Caroline Féry

The laryngeal effect in Korean: Phonology or phonetics? 171
Eon-Suk Ko

Time, tone and other things

The diphthong dynamics distinction in Swabian: How much timing is there in phonology? 195
Markus Hiller

Depression in Zulu: Tonal effects of segmental features 223
Philipp Strazny

Weakening processes in the Optimality framework 241
K. G. Vijayakrishnan

Base joint configuration in Sign Language of the Netherlands:
Phonetic variation and phonological specification **257**
Onno Crasborn and Els van der Kooij

Author index **289**

Language index **297**

Subject index **301**

Table of contents, volume II **305**

CURRENT ISSUES IN LINGUISTIC THEORY

Prof.dr. E.F.K. Koerner
Institut für Sprachwissenschaft, Universität zu Köln
D-50923 KÖLN, Germany
efk.koerner@uni-koeln.de

The *Current Issues in Linguistic Theory* (CILT) series is a theory-oriented series which welcomes contributions from scholars who have significant proposals to make towards the advancement of our understanding of language, its structure, functioning and development. CILT has been established in order to provide a forum for the presentation and discussion of linguistic opinions of scholars who do not necessarily accept the prevailing mode of thought in linguistic science. It offers an alternative outlet for meaningful contributions to the current linguistic debate, and furnishes the diversity of opinion which a healthy discipline must have. A complete list of titles in this series can be found on the publishers website, **www.benjamins.com/jbp**

200. ANTTILA, Raimo: *Greek and Indo-European Etymology in Action. Proto-Indo-European *aǵ-.* 2000.
201. DRESSLER, Wolfgang U., Oskar E. PFEIFFER, Markus PÖCHTRAGER and John R. RENNISON (eds.): *Morphological Analysis in Comparison.* 2000.
202. LECARME, Jacqueline, Jean LOWENSTAMM and Ur SHLONSKY (eds.): *Research in Afroasiatic Grammar. Papers from the Third conference on Afroasiatic Languages, Sophia Antipolis, 1996.* 2000.
203. NORRICK, Neal R.: *Conversational Narrative. Storytelling in everyday talk.* 2000.
204. DIRVEN, René, Bruce HAWKINS and Esra SANDIKCIOGLU (eds.): *Language and Ideology. Volume 1: cognitive theoretical approaches.* 2001.
205. DIRVEN, René, Roslyn M. FRANK and Cornelia ILIE (eds.): *Language and Ideology. Volume 2: descriptive cognitive approaches.* 2001.
206. FAWCETT, Robin P.: *A Theory of Syntax for Systemic Functional Linguistics.* 2000.
207. SANZ, Montserrat: *Events and Predication. A new approach to syntactic processing in English and Spanish.* 2000.
208. ROBINSON, Orrin W.: *Whose German? The* ach/ich *alternation and related phenomena in 'standard' and 'colloquial'.* 2001.
209. KING, Ruth: *The Lexical Basis of Grammatical Borrowing. A Prince Edward Island French case study.* 2000.
210. DWORKIN, Steven N. and Dieter WANNER (eds.): *New Approaches to Old Problems. Issues in Romance historical linguistics.* 2000.
211. ELŠÍK, Viktor and Yaron MATRAS (eds.): *Grammatical Relations in Romani. The Noun Phrase.* 2000.
212. REPETTI, Lori (ed.): *Phonological Theory and the Dialects of Italy.* 2000.
213. SORNICOLA, Rosanna, Erich POPPE and Ariel SHISHA-HALEVY (eds.): *Stability, Variation and Change of Word-Order Patterns over Time.* 2000.
214. WEIGAND, Edda and Marcelo DASCAL (eds.): *Negotiation and Power in Dialogic Interaction. 2001.*
215. BRINTON, Laurel J.: *Historical Linguistics 1999. Selected papers from the 14th International Conference on Historical Linguistics, Vancouver, 9-13 August 1999.* 2001..
216. CAMPS, Joaquim and Caroline R. WILTSHIRE (eds.): *Romance Syntax, Semantics and L2 Acquisition. Selected papers from the 30th Linguistic Symposium on Romance Languages, Gainesville, Florida, February 2000.* 2001.
217. WILTSHIRE, Caroline R. and Joaquim CAMPS (eds.): *Romance Phonology and Variation. Selected papers from the 30th Linguistic Symposium on Romance Languages, Gainesville, Florida, February 2000.* 2002.
218. BENDJABALLAH, S., W.U. DRESSLER, O. PFEIFFER and M. VOEIKOVA (eds.): *Morphology 2000. Selected papers from the 9th Morphology Meeting, Vienna, 24–28 February 2000.* 2002.

219. ANDERSEN, Henning (ed.): *Actualization. Linguistic Change in Progress.* 2001.
220. SATTERFIELD, Teresa, Christina TORTORA and Diana CRESTI (eds.): *Current Issues in Romance Languages. Selected papers from the 29th Linguistic Symposium on Romance Languages (LSRL), Ann Arbor, 8-11 April 1999.* 2002.
221. D'HULST, Yves, Johan ROORYCK and Jan SCHROTEN (eds.): *Romance Languages and Linguistic Theory 1999. Selected papers from 'Going Romance' 1999, Leiden, 9-11 December.* 2001.
222. HERSCHENSOHN, Julia, Enrique MALLÉN and Karen ZAGONA (eds.): *Features and Interfaces in Romance. Essays in honor of Heles Contreras. 2001.*
223. FANEGO, Teresa, María José LÓPEZ-COUSO and Javier PÉREZ-GUERRA (eds.): *English Historical Syntax and Morphology. Selected papers from 11 ICEHL, Santiago de Compostela, 7-11 September 2000.* 2002.
224. FANEGO, Teresa, Belén MÉNDEZ-NAYA and Elena SEOANE (eds.): *Sounds, Words, Texts and Change. Selected papers from 11 ICEHL, Santiago de Compostela, 7-11 September 2000.* 2002.
225. SHAHIN, Kimary N.: *Postvelar Harmony.* 2002.
226. LEVIN, Saul: *Semitic and Indo-European. Volume II: comparative morphology, syntax and phonetics; with observations on Afro-Asiatic.* 2002.
227. FAVA, Elisabetta (ed.): *Clinical Linguistics. Theory and applications in speech pathology and therapy.* 2002.
228. NEVIN, Bruce E. (ed.): *The Legacy of Zellig Harris. Language and information into the 21st century. Volume 1: philosophy of science, syntax and semantics.* n.y.p.
229. NEVIN, Bruce E. and Stephen JOHNSON (eds.): *The Legacy of Zellig Harris. Language and information into the 21st century. Volume 2: computability of language and computer applications.* 2002.
230. PARKINSON, Dilworth B. and Elabbas BENMAMOUN (eds.): *Perspectives on Arabic Linguistics XIII-XIV. Papers from the Thirteenth and Fourteenth Annual Symposia on Arabice Linguistics.* 2002.
231. CRAVENS, Thomas D.: *Comparative Historical Dialectology. Italo-Romance clues to Ibero-Romance sound change.* 2002.
232. BEYSSADE, Claire, Reineke BOK-BENNEMA, Frank DRIJKONINGEN and Paola MONACHESI (eds.): *Romance Languages and Linguistic Theory 2000. Selected papers from 'Going Romance' 2000, Utrecht, 30 November - 2 December.* 2002.
233. WEIJER, Jeroen van de, Vincent J. van HEUVEN and Harry van der HULST (eds.): *The Phonological Spectrum. Part I: Segmental structure.* 2003.
234. WEIJER, Jeroen van de, Vincent J. van HEUVEN and Harry van der HULST (eds.): *The Phonological Spectrum. Part II: Suprasegmental structure.* 2003.
235. LINN, Andrew R. and Nicola MCLELLAND (eds): *Standardization. Studies from the Germanic languages.* 2002.
236. SIMON-VANDENBERGEN, Anne-Marie, Miriam TAVERNIERS and Louise RAVELLI: *Grammatical Metaphor. Views from systematic functional linguistics.* n.y.p.
237. BLAKE, Barry J. and Kate BURRIDGE (eds.): *Historical Linguistics 2001.Selected papers from the 15th International Conference on Historical Linguistics, Melbourne, 13–17 August 2001.* n.y.p.
238. NÚÑEZ-CEDENO, Rafael, Luis LÓPEZ and Richard CAMERON (eds.): *A Romance Perspective on Language Knowledge and Use. Selected papers from the 31st Linguistic Symposium on Romance Languages (LRSL), Chicago, 19–22 April 2001.* n.y.p.
239. ANDERSEN, Henning (ed.): *Language Contacts in Prehistory. Studies in Stratigraphy. Papers from the Workshop on Linguistic Stratigraphy and Prehistory at the Fifteenth International Conference on Historical Linguistics, Melbourne, 17 August 2001.* n.y.p.
240. JANSE, Mark and Sijmen TOL (eds.): *Language Death and Language Maintenance. Theoretical, practical and descriptive approaches.* n.y.p.