Input-based Phonological Acquisition

Tania S. Zamuner

OUTSTANDING DISSERTATIONS IN LINGUISTICS

Edited by

Laurence Horn
Yale University

A ROUTLEDGE SERIES

OUTSTANDING DISSERTATIONS IN LINGUISTICS

LAURENCE HORN, *General Editor*

TURN-TAKING IN ENGLISH AND JAPANESE
Projectability in Grammar, Intonation and Semantics
Hiroko Furo

MORPHOLOGICALLY GOVERNED ACCENT IN OPTIMALITY THEORY
John Alderete

MINIMAL INDIRECT REFERENCE
A Theory of the Syntax-Phonology Interface
Amanda Seidl

DISTINCTIVENESS, COERCION AND SONORITY
A Unified Theory of Weight
Bruce Morén

PHONETIC AND PHONOLOGICAL ASPECTS OF GEMINATE TIMING
William H.Ham

VOWEL REDUCTION IN OPTIMALITY THEORY
Katherine Crosswhite

AN EFFORT BASED APPROACH TO CONSONANT LENITION
Robert Kirchner

THE SYNCHRONIC AND DIACHRONIC PHONOLOGY OF EJECTIVES
Paul D.Fallon

GRAMMATICAL FEATURES AND THE ACQUISITION OF REFERENCE
A Comparative Study of Dutch and Spanish
Sergio Baauw

AUDITORY REPRESENTATIONS IN PHONOLOGY
Edward S. Flemming

THE TYPOLOGY OF PARTS OF SPEECH SYSTEMS
The Markedness of Adjectives
David Beck

THE EFFECTS OF PROSODY ON ARTICULATION IN ENGLISH
Taehong Cho

PARALLELISM AND PROSODY IN THE PROCESSING OF ELLIPSIS SENTENCES
Katy Carlson

PRODUCTION, PERCEPTION, AND EMERGENT PHONOTACTIC PATTERNS
A Case of Contrastive Palatalization
Alexei Kochetov

RADDOPPIAMENTO SINTATTICO IN ITALIAN

A Synchronic and Diachronic Cross-Dialectical Study
Doris Borrelli

PRESUPPOSITION AND DISCOURSE FUNCTIONS OF THE JAPANESE PARTICLE
Mo
Sachiko Shudo

THE SYNTAX OF POSSESSION IN JAPANESE
Takae Tsujioka

COMPENSATORY LENGTHENING
Phonetics, Phonology, Diachrony
Darya Kavitskaya

THE EFFECTS OF DURATION AND SONORITY ON CONTOUR TONE DISTRIBUTION
A Typological Survey and Formal Analysis
Jie Zhang

EXISTENTIAL FAITHFULNESS
A Study of Reduplicative TETU, Feature Movement, and Dissimilation
Caro Struijke

PRONOUNS AND WORD ORDER IN OLD ENGLISH
With Particular Reference to the Indefinite Pronoun Man
Linda van Bergen

ELLIPSIS AND WA-MRKING IN JAPANESE CONVERSATION
John Fry

WORKING MEMORY IN SENTENCE COMPREHENSION
Processing Hindi Center Embeddings
Shravan Vasishth

INPUT-BASED PHONOLOGICAL ACQUISITION

Tania S.Zamuner

Routledge
New York & London

Published in 2003 by
Roudedge
29 West 35th Street
New York, NY 10001
www.routledge-ny.com

Published in Great Britain by
Routledge
11 New Fetter Lane
London EC4P 4EE
www.routledge.co.uk

Routledge is an imprint of the Taylor & Francis Group

This edition published in the Taylor & Francis e-Library, 2005.

"To purchase your own copy of this or any of Taylor & Francis or Routledge's collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk."

Library of Congress Cataloging-in-Publication Data

Zamuner, Tania S.
Input-based phonological acquisition/by Tania S.Zamuner.
p. cm. —(Outstanding dissertations in linguistics)
Includes bibliographical references and index.
ISBN 0-415-96789-9 (hb)
1. Language acquisition. 2. Grammar, Comparative and general—Phonology. I. Title.
II. Series.
P118.Z358 2003
401′.93–dc21 2003046798

ISBN 0-203-00990-8 Master e-book ISBN

Table of Contents

Acknowledgements

This research was supported by a Cognitive Science (Language) Graduate Fellowship from the Natural Sciences and Engineering Research Council of Canada.

I was fortunate to have joint advisors of LouAnn Gerken and Michael Hammond who have guided and fed me, read numerous drafts, given me invaluable and endless advice, insight and encouragement, and helped me see through to the end. A special thanks to the third member of my committee, Diane Ohala, for her perspicacious comments, especially for insights in analysing the child language data and lastly for providing the pictures used in Experiment 2.

I have benefited greatly from useful discussions in LouAnn Gerken's Acquisition Lab and from help in all aspects of conducting experiments. Its members include(d): Allyson Carter, Ahnie Dysinger, Susan Garcia, Janet Gatewood, Rebecca Gomez, Andrea Massar, Jessica Maye, Tanya Noth, Erika Nurmsoo, Asami Nozuki, Diane Ohala, Maryam Sadrzadeh, and Rachel Wilson. Thanks also to Asami Nozuki and Maryam Sadrzadeh for help in running and coding the experiments. Thanks to Marlena Creusere for proofreading. Other people have been influential in my linguistics enthusiasm and career especially Heather Goad, Diana Archangeli, Dick Demers, Nigel Duffield, Cecile McKee, Laura Pettito, and Glyne Piggott. And thanks to Weibke Brockhaus, Taehong Cho, Suzanne Curtin, Paul Luce, Ian Maddieson, Janet Pierrehumbert, and Elena Plante for helpful discussions.

I'm grateful for the academic support and esprit de corps from my fellow students in linguistics, especially Meg O'Donnell, Feng Wei, Debbie Cole, and Robert Kennedy. In administration—in linguistics, Jennie Bradley, Gwen Webber and in particular Rosemary Emery, in Speech and Hearing, Pam Adams, Julie Mills and Donna Douglas, exercised admirable patience and made the departments agreeable places to be.

My friends made Tucson a home, notably Steve Hamann and Frankie Ridolfsky who knocked on my door the first day I arrived, when I knew no one and had no furniture. Sierk Pötting shared 4:30 am coffees with me on the patio while people went for their morning jogs. John Lapeyre helped me copiously on the programming needed throughout. And thanks to others from afar: Keith Daniel, Ann Brodie-Brown, Bella Malavolta, Chuck E. Marchant, Norval Smith, Dylan MacKay and Jennifer Bird.

My family has provided me with a constant supply of love and unfailing support. This dissertation is surreptitiously dedicated to them—to the Zalerota Welding Gun.

And thank you to Erik Jan van der Torre.

Preface

A few changes have been made from the dissertation submitted to the University of Arizona in July 2001. Notably, the name for the input-based hypothesis for language acquisition was changed from the General Pattern Learning Hypothesis (GPLH) to the Specific Language Grammar Hypothesis (SLGH). The reason for this change was that the former seems to make predictions about other cognitive abilities, while the latter makes predictions specifically about language. Given that this dissertation is about the acquisition of phonological structures, it seems more appropriate to refer to the input-based hypothesis for language acquisition as the SLGH. The SLGH hypothesis does not exclude the possibility that language is acquired as other cognitive abilities are, but given the fact that I do not address this issue, I've framed the hypothesis in its slightly more conservative form. I have also revised the prediction based on Experiment 1 to exclude coda consonants that occurred in words not spontaneously produced by children. This gives a more accurate prediction because it takes into account the possibility that some of the words in Experiment 1 were not very familiar to children and therefore not good examples from which to base children's productions of codas. All subsequent analyses reflect this change. In fact, this change made no difference in the pattern of results. Lastly, I have simplified the analyses in Experiment 2, and made a more detailed coding of children's responses.

Abstract

This dissertation contrasts two theories of language acquisition. The first theory, which has been dominant in generative linguistics, argues that acquisition is primarily mediated by innate properties of language provided by universal grammar (Universal Grammar Hypothesis—UGH). Data to support this theory are the parallels between cross-linguistic data and child language. However, the structures that are frequent across language are also typically the most frequent within a specific language. This confounding of cross-linguistic and language specific data is consistent with the view that language is acquired based on the patterns in the ambient language or input (Specific Language Grammar Hypothesis—SLGH). These theories are contrasted by examining children's acquisition of coda consonants in CVC words.

The UGH of coda preferences was based on previous research and on frequency analyses of codas in CVC words from 35 languages. Results showed that languages prefer coronal codas and sonorant codas. These cross-linguistic preferences are interpreted as reflecting UG. Predictions of the SLGH were established through an examination of English codas in CVC words from a number of different sources. This revealed the frequency of codas in the input, upon which the SLGH was based. In order to determine which hypothesis better predicts children's coda acquisition, data were then collected from previously published research, from CHILDES, and from an experiment designed to test children's production of English codas.

To evaluate the UGH, children's coda productions were analysed to determine whether the preferred codas were coronals or sonorants. Results did not show that these codas were favoured. To evaluate the SLGH, analyses determined whether there were significant correlations between children's coda productions and the frequency of English codas. Results showed that these relationships were significant.

The role of the input was further examined in an experiment designed to test children's productions of the same coda in non-words controlled for phonotactic probabilities. Results showed that phonotactic probabilities played a significant role in accounting for children's production of the same coda in different words.

The results support an input-based account of phonological acquisition. Thus, language acquisition is best characterized with respect to patterns in the ambient language, where frequently occurring properties of the input serve to organize children's linguistic representations. The research here illustrates the importance of considering the input in children's acquisition of phonological structures.

INPUT-BASED PHONOLOGICAL ACQUISITION

CHAPTER 1
Accounts of acquisition
Universal Grammar and the input

"The Simpsons"

Herb: It's a "Baby Translator." It measures the pitch, the frequency and the urgency of a baby's cries. Then it tells whoever's around, in plain English, exactly what the baby's trying to say. Everything from "Change me," to "Turn off that damn Raffi record."

* * * * * *

Herb: Maggie, who brought me my fortune. Just name anything you want in this world.
Maggie: Blah blah blah.
Translator: I want what the dog's eating.
Homer: D'oh!

1.1.
Focus of research

The invention of a baby translator presupposes that children make systematic errors in their speech that reflect a system of organisation. Although this dissertation does not provide a blueprint for a baby translator, it does ask how children's productions can help us understand how children interact with the ambient language, and it thereby explores the mechanisms that underlie children's ability to successfully acquire language.

The primary question asked in this dissertation concerns the nature of language acquisition. In particular, I contrast two views. One view, which has been dominant in generative linguistics, is that language acquisition is best characterised as the unfolding of highly specified innate abilities (Chomsky, 1981). On this view, children come to the language-learning task with innate knowledge of how language is structured, and the input serves only to trigger innately specified options. The opposing view is that language is acquired based on the patterns in the ambient language or the input (Olmsted, 1966, 1971). On this account, language acquisition is characterised as a domain-neutral learning process in which frequently occurring properties of the input serve to organise children's linguistic representations. I contrast these theories by examining young children's acquisition of coda consonants.

This chapter focuses on nature versus nurture accounts for the observation that children's productions mirror cross-linguistic markedness (to be defined later), and describes the case employed here to test these hypotheses. To begin, I describe the noted parallel between cross-linguistic markedness and child language. The primary account of this relationship relies on the Chomskian notion of innate universal grammar (UG). I define this theory of grammar acquisition as the Universal Grammar Hypothesis (UGH). I demonstrate that the data taken to support this account are often confounded, given that the distributions of sound patterns in a single language often encode what is argued to be innately provided by UG. I further describe a subset of research that illustrates children are sensitive to the frequency of sound patterns in the input language. The view that children acquire language according to the input is introduced as an alternative to the UGH, and defined as the Specific Language Grammar Hypothesis (SLGH). The acquisition of coda consonants by children acquiring English is presented as a means for distinguishing between the two hypotheses. Subsequently, the latter half of Chapter 1 introduces this dissertation's focus: the test case of word-final codas, which allows for the examination of the relative roles of innate knowledge versus input frequency in phonological acquisition.

1.2. Universal Grammar

One problem in discussing UG is that there is little clarity in the term's use and meaning. Moreover, the definition and interpretation of "universal" and "universal grammar" differs across researchers and across the sub-fields of linguistics, such as in phonology and syntax. The following discussion will centre on two different characterisations of UG.

1.2.1. Establishing Universal Grammar

There are two central interpretations of UG. The first interpretation is that UG contains the set of properties present in all languages. This set comprises unrestricted or absolute universals. For example, all languages have CV syllables (Spencer, 1996, p.82). Unrestricted universals refer exclusively to the necessary properties that define all languages or what are also called the "unmarked" properties of languages. This is formally represented in (1) (adapted from Hammond, Moravcsik & Wirth, 1988), and defined as: for all languages, Y.

(1)	Unrestricted universal Y CV syllables	
Type 1 language	+	occurs
Type 2 language	−	*does not occur*

All languages have CV syllables (Type 1); moreover, there are no languages without CV syllables (Type 2). Unrestricted universals are by definition based on cross-linguistic research.

The second interpretation of UG is that it defines the limits of human language and the extent to which languages can vary. Thus, UG contains and is defined by what is unmarked and marked in language. The term "unmarked" refers to the properties of language that are common and frequent (which can be different from what is frequent in any one language), whereas "marked" defines properties that are less common and less frequent. Various methods are used to determine the extent to which languages can vary, or to determine what is unmarked and marked in languages. This can be based on restricted universals, language processes, language change, patterns seen in child language and aphasia, and on the distribution of sounds in languages, all of which are discussed below. Moreover, the cross-linguistic markedness factors on which UG is defined differ according to different theoretical perspectives. For the purposes here, patterns established by all methods are taken as evidence for cross-linguistic markedness, and in turn, as evidence for UG.

Some argue that markedness is necessarily based solely on restricted universals (also called implicational universals or typological universals). These are defined as: for all languages, if X, then Y (Greenberg, 1963). For instance, if a language possesses obstruent codas, then it possesses sonorant codas. The logic is schematised in (2) (adapted from Hammond et al. 1988).

(2) Restricted universal

	X	→	Y	
	obstruent codas		sonorant codas	
Type 1	+		+	occurs
Type 2	−		−	occurs
Type 3	−		+	occurs
Type 4	+		−	*does not occur*

In other words, the presence of obstruent codas entails the presence of sonorant codas; no languages exist with only obstruent codas (Type 4 languages do not occur). The unmarked property is implied by the marked one (Greenberg, 1966); thus, sonorant codas are unmarked and obstruent codas are marked.

Markedness claims are also based on language processes and language change (Greenberg, 1974). For example, in neutralisation, one feature will be neutralised more often than the other feature, such as in final devoicing (Greenberg, 1966). In this case, laryngeal features are restricted in coda position (see Lombardi, 1991). In Catalan (3) and Turkish (4), voiced obstruents become voiceless in word-final position. For example, in Catalan, the /b/ in /ʎobə/ becomes [p] word-finally.

(3) Catalan (Hualde, 1992)

ʎobə	*wolf (fem)*	ʎop	*wolf (masc.)*
əmigə	*friend (fem)*	əmik	*friend (masc.)*
frənsɛzə	*French (fem)*	frənsɛs	*French (masc.)*

(4) Turkish (Lewis, 1967)

tʃeridi	*tape (acc)*	tʃerit	*tape*
dibi	*bottom (acc)*	dip	*bottom*
grubu	*group (acc)*	grup	*group*

The feature that is most likely to be neutralised is argued to be marked, and the resulting feature is unmarked. With syllable-final devoicing, voiced stops are marked, subsequently becoming voiceless, which are unmarked.

Markedness claims are also based on patterns seen in child language acquisition and pathology, where children and aphasic adults are argued to produce the least marked structures in language (Blumstein, 1973; Jakobson, 1941/1968). Children and aphasics delete codas to produce CV syllables, which are also the least marked syllables cross-linguistically. This adds circularity to arguments of language acquisition that are based on UG.

Lastly, there are statistical, substantive or distributional universals (Greenberg, 1954, 1966, 1974; Hammond et al. 1988; Trubetzkoy, 1939/1969; Zipf, 1935/1965). These types of generalisations are not absolute, but rather, they describe languages' tendencies to pattern in a particular way. This originated with the Prague school of linguistics, which first identified the relationship between markedness and frequency in phonology (Trubetzkoy, 1931, cited in Greenberg, 1966, p.11). This school stressed the importance of statistical information in language from a variety of sources, such as dictionaries and texts. The rationale for considering frequency with respect to UG is that frequency is generally related to markedness across languages (Greenberg, 1974), and markedness is translated as defining the properties of UG. Not only is markedness connected to frequency across languages, but it is also encoded in the frequency of sound patterns within a single language. (Frequency has also been argued to be not formally related to markedness, Leuninger, 1983, p.35.)

To summarise, there are two primary interpretations of UG. The first is based on unrestricted universals and refers solely to the unmarked properties of language. The second definition for UG is that it defines the possible limits of human language and refers to both the unmarked and marked properties of language. Markedness is established by a variety of means, such as by restricted universals and frequency. The characterisation of UG that is relevant for language acquisition is subject to debate; however, constant across all theories is that UG minimally provides to children the unmarked properties or settings for language, as part of children's innate genetic endowment.

1.2.2. Universal Grammar and child language: The parallel between cross-linguistic markedness and child language

The view that language acquisition is mediated by UG stems from Jakobson's theory of acquisition, within the framework of the Prague School of Linguistics. Jakobson's theory illustrates how adult language patterns can be extended to acquisition. For Jakobson, acquisition is a process of acquiring contrasts in meaningful productions. In this sense he makes a distinction between babbling and communication—word use marks the beginning of the development of a phonological system. Children's acquisition of contrasts is governed by the same "general laws of irreversible solidarity" that govern adult language or the laws that govern the phonemic systems of languages. Thus, Jakobson makes the important observation that there are similarities between child language and adult language. He further notes that the order of acquisition can be predicted from cross-linguistic implicational universals. This order is invariant across children arguably because the knowledge is innate, however, the speed at which children acquire contrasts can vary. Roughly, Jakobson predicts that the first opposition to be acquired is between a consonant and a vowel. From there, finer distinctions are made and further contrasts are built.

To illustrate the parallel between adult language and child language, take the example of syllable structure. Languages of the world exhibit certain universal patterns. For example, the universally unmarked syllable shape is CV, such that all languages permit CV syllables (Blevins, 1995; Clements & Keyser, 1983; Greenberg, 1978). The least marked languages permit only this syllable type (CV) and do not permit coda consonants, e.g., Arabela, Cayuvava (5), Hawaiian (6), Hua, Kikuyu, Mazateco, Sesotho, Swahili, Tongan and Twi.

(5)	Cayuvava (Key, 1967)
bore	*food*
ni	*eat*
korotʃo	*black insect*
mūrūtū	*tiny bee*

(6)	Hawaiian (Walsh & Biggs, 1966)
ʔele	*black*

ʔula	*red*
hale	*house*
hele	*octopus*

In these languages, coda consonants are not permitted, consequently all syllables are of a CV (or V) shape. In Hawaiian this effect is mirrored in English loan words. Words with syllables that do not conform to CV syllables shapes are modified (7).

(7) Hawaiian English loan words Elbert & Pukui, 1979, cited in Spencer, 1996)

English	Hawaiian	
kroke	koloke:	*croquet*
wɪski	wekeke:	*whiskey*

For example, the English word "croquet" begins with two consecutive consonants /kr/. In Hawaiian, which does not allow consonant clusters, /kr/ is modified through epenthesis of [o] between the illicit /kr/ sequence, producing [koloke:]. (The change of /r/ to [l] results from other processes.)

Interestingly, children's early word productions are largely characterised by this same open syllable shape. Young children initially delete or modify their productions of codas or other marked syllable structures to conform to CV shapes as in (8) and (9). For example, a child attempting to produce the final /n/ in "man," might delete /n/ and produce [mæ], [mɛ], or[mʌ].

(8) English (Demuth & Fee, 1995)

Child	Adult	
[mæ], [mɛ], [mʌ]	/mæn/	*man*
[bo], [bɑ], [bo:]	/bɑl/	*ball*

(9) Dutch (Fikkert, 1994)

Child	Adult	
[ka:], [kɑ]	/kla:r/	*ready*
[pu:]	/pu:s/	*(pussy) cat*
[ka:][kɔ]	/tɔk/	*chicken sound*
[pa:], [pɑ]	/pa:rt/	*horse*
[ta:], [tɑ]	/dɑt/	*that*

From these examples, we can see that children initially delete coda consonants; thus, their productions conform to an onset-peak or CV shape. Another strategy children use to avoid producing codas is to epenthesize a vowel after final consonants, e.g., CVC →CVCV, /tʃɪk/ 'chick' is produced as [dɛ́di] (PJ 23 months, taken from Demuth & Fee, 1995). Also see Fee and Ingram (1982). One interpretation of this phenomenon is that final consonant deletion results from innate pressures to conform words to the preferred CV shape. Thus, these errors have been traditionally interpreted as a reflection of innate UG (e.g., Demuth & Fee, 1995; Faingold, 1990; Jakobson, 1941/1968; Levelt, Schiller & Levelt, 2000; Levelt & Vijver, in press; Mowrer & Burger, 1991; Ohala, 1996, Vihman & Ferguson, 1987 and others).[1]

The observation that children's productions mirror cross-linguistic markedness has fuelled the hypothesis that substantial parts of linguistic structure are innate and that language acquisition is mediated by UG. The standard linguistic argument is that the parallel between cross-linguistic markedness and child language is the effect of UG on both systems.

1.2.3. Universal Grammar Hypothesis of language acquisition

Within generative theories of linguistics, the unmarked properties of language or the core grammar (Chomsky, 1981) comprise children's initial hypotheses about language. Thus, children's initial CV productions are explained by appealing to the CV syllable as the unmarked syllable type or state. As children's grammars develop, they move away from the core grammar towards a language-specific grammar. Consequently, children produce more marked syllable structures.

Although markedness and UG are hypothesised to motivate patterns seen in child language, patterns in acquisition do not always follow UG. Children do not necessarily initially produce what is unmarked and acquire marked structures last (Hyams, 1986). For instance, although /l/ and /r/ are preferred consonants cross-linguistically, these sounds are difficult for children to perceive and produce (Poole, 1934; Welleman, Case, Mengert & Bradbury, 1931). Given that children exhibit perceptual and production constraints that differ from adults (as with /l/ and /r/), this raises an interesting question of how to reconcile these differences, because factors other than UG are evidenced in acquisition, and domains can interact with one another. These are complex issues that need to be addressed. As a starting point, however, I take the standard approach assumed in the acquisition literature, where adult language patterns are seen as reflecting the properties of UG.

In summary, the hypothesis that language acquisition is guided by innate principles of language structure, which are provided by UG, is defined as the Universal Grammar Hypothesis (UGH) (10):

(10) Universal Grammar Hypothesis (UGH): Children acquire language based on innate properties of language, which guide children's acquisition of phonology.

Recall that UG is evidenced through cross-linguistic markedness effects. Thus, parallels between child language and cross-linguistic markedness can be represented as the product of UG. Because UG is innate, it will be represented in all human languages (as seen in cross-linguistic markedness). Similarly, UG is innately provided to children; thus, children's productions reflect this genetic knowledge. Interestingly, markedness is seen as both as evidence for UG and as the product of UG.

---

[1] The main justification for innate language is the induction problem in acquisition, which argues that the learner's hypothesis space needs to be constrained to prevent the learner from positing an overly general grammar. If the child assumes a larger grammar than that of their language community, they will need access to negative evidence to restrict their grammar to the correct one. It is claimed, however, that children do not have access to negative evidence during acquisition. Thus, it is assumed that children come equipped with knowledge about how language is structured, and that this knowledge constrains the initial hypotheses children make about language. Most of the induction problem has centred on syntactic acquisition; however some researchers in phonological acquisition have also drawn on these arguments (e.g., Dell, 1981; Goad & Brannen, in press; Ohala, 1992). UG is proposed as a solution to the induction problem.

1.3.
The input

This section centres on the alternative to the UG hypothesis for language acquisition, which is that language acquisition is best described with respect to the patterns in the input or ambient language. This hypothesis is in part motivated by the large body of research that illustrates children are sensitive to the frequency of patterns in the ambient language.

1.3.1.
Child language reflects the language specific input

There is considerable evidence that child language reflects the language-specific input at an early age. For example, nine-month-old infants show a preference for their language's more common sound patterns (Jusczyk, Luce & Charles-Luce, 1994), and around 10 months, infants begin to ignore phonetic contrasts that are not phonemic in the language they are acquiring (Werker & Tees, 1984). This sensitivity to the input is also evident in young children's pre-linguistic vocalisations, where infants' babbling reflects the ambient language (see review by Vihman & Boysson-Bardies, 1994).

The effects of the input or the ambient language are further mirrored in children's linguistic productions, at both the segmental and syllable structure levels. For example, high frequency sounds in the ambient language are more likely to be produced by children than are low frequency sounds. The palato-alveolar affricate /tʃ/, which has a low frequency in English, is acquired late; however, in Spanish, /tʃ/ has a high frequency and is acquired earlier (Macken, 1995). Similarly, results from Pye, Ingram and List (1987) showed that children's initial consonant productions in Quiché and English were significantly correlated with the frequency of initial consonants in their respective languages. As in Spanish, the affricate /tʃ/ is acquired early in Quiché, and the early acquisition of /tʃ/ is argued to be linked to its frequency in child directed speech. In Finnish /d/ is acquired late, unlike in other languages, and this is hypothesised to reflect the phoneme's low frequency in Finnish (Macken, 1995). Word initial /v/ is acquired earlier in Spanish, Estonian, and Bulgarian than in English, arguably because of the phoneme's high frequency in the former languages (Ingram, 1988).

At the prosodic level, Levelt, et al. (2000) and Levelt and Vijver's (in press) examinations of Dutch children's acquisition of syllable types illustrate that children who learn languages with a complex or marked syllable structure must choose how to build syllable structure beyond the initial core CV syllables (arguably children choose between different learning paths given by UG). Faced with choice, children first acquire the syllable structure that is most frequent in Dutch. Another example, from the acquisition of coda consonants, demonstrates that children learning English produce codas before children learning Spanish (Demuth, 2001; Gennari & Demuth, 1997). It is suggested that this can be attributed to differences in coda frequencies in the two languages; codas are more frequent in English than in Spanish (Roark & Demuth, 2000).

Frequency has also been shown to be a factor in children's early word productions, where children's first words are often the most frequently presented words in the input (Harris, Barrett, Jones & Brookes, 1988; Naigles & Hoff-Ginsberg, 1998); however, this relationship becomes less pronounced as children's vocabulary sizes grow (Barrett, Harris & Chasin, 1991; Hart, 1991). One can also see the effect of the frequency of word category in children's productions across languages. For example, children produce determiners in Spanish earlier than in German. This is in part due to differences in prosody, where determiners in Spanish participate in more frequent prosodic patterns (Lleó & Demuth, 1998).

1.3.2. Specific Language Grammar Hypothesis of language acquisition

Children's early sensitivity to the frequency of sound patterns in the ambient language suggests an alternative to the UGH for acquisition. Given that cross-linguistic markedness (upon which the present version of UG is based) is reflected in a language's distribution of sounds, it follows that the properties of language that are argued to be innate are often the most frequent properties in any one language. As proposed earlier, one could argue that children initially produce unmarked structures due to innate UG; however, one could equally argue that children initially produce these structures because they are the most frequent patterns in their ambient language.

Thus, an alternative to the innate universal account of acquisition is the Specific Language Grammar Hypothesis, which states that language acquisition is guided by the frequency of sound patterns in the ambient language (11):

(11) Specific Language Grammar Hypothesis (SLGH): Children acquire language based on the patterns in the input language, which drive children's acquisition of phonology.

In this sense, frequently occurring patterns in the input serve to organise children's linguistic representations. The prediction based on this hypothesis is that children should initially produce the more frequently occurring patterns and structures in their ambient language, before producing the less frequent ones. To evaluate this hypothesis, children's productions are compared to the language-specific input to determine whether there are significant correlations between the two. Note that even a frequency-based theory requires innate mechanisms to focus the language learner on the initial information needed for language acquisition (Morgan, 1990); however, the SLGH differs from the UGH in that innate *linguistic* knowledge is not required.

The type of information accounted for by the two approaches is often quite different. Innate theories of language acquisition centre on formal accounts of how children acquire the language structure of a specific language, e.g., arguing how children determine the ordering of innate constraints with respect to one another. General learning models tend to focus on how children learn the distribution of sounds in their ambient language, such as learning which sounds are permitted in different prosodic positions. Differences both within and across the different theories approaches to acquisition make it difficult to directly compare a UG based theory of acquisition to a more general learning theory of acquisition. This raises questions about what unit or level of analyses the hypotheses' predictions should be based on, in order to allow for a reasonable comparison between the two theories. This question will be revisited in §3.5.

1.4. The research question

The research focuses on the often-noted similarities between child language and cross-linguistic data, and explores the role of frequency in the ambient language. The argument is *not* that child language is the result of the input alone. I consider a detailed investigation of the role of frequency in children's acquisition of phonology in order to determine whether the UGH or the SLGH better accounts for the acquisition of codas. Regardless of whether one argues for a UG or an input-based account of acquisition, both approaches need to be considered as they often make the same predictions. Both theories regularly converge on the same information, yet apply opposite interpretations and frameworks to the data. The two theories are not necessarily mutually exclusive, however, the input account has not been as thoroughly explored.

These issues fit into a larger research agenda in which there are a number of logical possibilities for the interaction of these two domains with respect to acquisition. In the table below, possible language acquisition patterns are laid out with respect to UG (as evidenced by adult cross-linguistic markedness), the Language Specific Patterns (the ambient language to which children are exposed), and Child Language Patterns (young children's productions).

Table 1 Logical possibilities for language acquisition

	UG (X-Linguistic data)	Language Specific Patterns (LSP)	Child Language Patterns (CLP)
a.	same	same	same
→b	same	different	same
→c.	different	same	same
d.	same	same	different
e.	different	different	different

The initial case in (a) is where the same pattern holds across UG, LSP and CLP. In this case, one cannot determine whether children are sensitive to innate UG or to the patterns in the ambient language (LSP). The following cases of (b) and (c) are where children's sensitivity to UG or LSP is disassociated. In (b) children produce the language-specific patterns over UG, and vice versa in (c). The cases of (d) and (e) are where children's patterns depart from UG and the input. For example, (d) and (e) might reflect children's cognitive or physical limitations where children have difficulties perceiving or producing /l/ and /r/, and produce these segments as vowels, despite the fact that these sounds are cross-linguistically unmarked and frequent. Further research questions are whether children's abilities and sensitivities to UG and LSP remain constant or change over time, and whether there are different patterns and/or interactions seen across different language domains. The research presented here explores the cases of (b) and (c) with respect to children's acquisition of coda consonants.

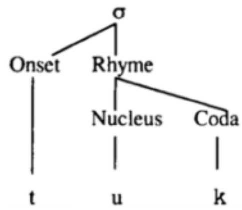
1.4.1.
Test case: The acquisition of coda consonants

A study of coda consonant acquisition offers a way of examining the interactions between the domains of cross-linguistic markedness (UGH), the distribution of sound patterns in a specific language (SLGH) and child language. This is because the distribution of codas is restricted both across languages and within languages, and as children's initial productions of coda consonants are also restricted. Cross-linguistically, some languages do not permit codas, and when they are permitted, languages favour codas with coronal place, voiceless laryngeal features and increased sonority. Different patterns, however, may hold in a single language (this will be further established in Chapters 2 and 3). Parallel to the cross-linguistic data, children initially delete and modify codas in ways that reflect the markedness of place, voice and in some instances, sonority. Although languages also prefer codas with voiceless laryngeal features, I will not consider this factor here. Research has shown that children produce laryngeal distinctions that are not initially audible to adults, even as late as 4; 6 (Macken & Barton, 1980; Scobbie, Gibbon, Hardcastle & Fletcher, 2000; Smith, 1979). A study of laryngeal features in coda position would thereby require acoustic measurements of children's productions and is not pursued here. The analyses in Chapter 2–5 will establish these cross-linguistic preferences, distribution of patterns in English, and child language acquisition facts. The goal will be to determine whether children's coda productions reflect cross-linguistic markedness and/or the language-specific input. Before laying out the analyses, I turn to a definition of coda consonants, place of articulation and sonority, and I further identify the assumptions taken and the criteria used in subsequent analyses.

1.4.1.1. *Definition of coda consonants*

A coda is defined as a syllable- or word-final consonant, such as the "k" in "toque" (12).[2]

(12)



The data considered throughout the dissertation are restricted to word-final codas in monosyllabic CVC words for two primary reasons. First, there is no consensus as to how intervocalic consonants are syllabified; different linguistic theories predict different syllabifications for medial consonants (e.g., Hammond, 1997; Kahn, 1976; Selkirk, 1982). In addition, native speaker's intuitions of how these consonants are syllabified are affected by vowel quality (lax versus tense), consonant type (liquids versus nasals versus obstruents), stress (stressed or unstressed), spelling (single or double letter) and morphology (monomorphemic versus bimorphemic) (e.g., Derwing, 1992; Fallows, 1981; Treiman & Danis, 1988; Zamuner & Ohala, 1999). Thus, determining the syllabic constituency of medial consonants is a complicated matter that would only confuse the issue at hand.

Second, the set of possible codas, the frequency of codas and children's production of codas are all affected by prosodic environment (Franke, 1912; Hammond, 1999; Zamuner, 1998a, 1998b; Zamuner & Gerken 1997). For example, Zamuner and Gerken found that 2-year-old children's productions of codas were dependent on prosodic environment. More specifically, both sonorants and obstruents were produced in stressed syllables, whereas in unstressed syllables children produced mainly sonorants. Again, this adds another complication, albeit an interesting one, to the question being considered. To circumvent these problems, the analyses are restricted codas to final position in monosyllabic CVC words.

1.4.1.2. *Representation of final consonants*

Although word-final consonants are generally syllabified as codas, not all phonological theories adopt this syllable structure representation. For example, it has been proposed that word-final consonants are syllabified as onsets of empty-headed syllables, rather than as codas (see for example, Kaye, Lowenstamm

---

[2] The goal of this dissertation is not to qualify or define a specific prosodic representation. I offer no arguments as to whether the coda is a constituent, but rather, I illustrate that there are consistent cross-language preferences in word-final position.

& Vergnaud, 1990; Harris, 1997; Piggott, 1991, 1999). Under this assumption, Goad (1997a) and Goad and Brannen (to appear) argue that children's first productions of final consonants are initially syllabified as onsets to account for the phonetic properties of these consonants (aspiration, length and homorganic nasal release). Regardless of the theory of representation that one adopts for these final-consonants, both theories predict that the same types of consonants will be preferred in word-final position.

Another assumption about the representation of word-final consonants in this dissertation is that these consonants are not syllabified as part of an appendix and that these consonants are not extraprosodic (Booij & Rubach, 1984; Borowsky, 1986; Halle & Vergnaud, 1980; Itô, 1986; McCarthy, 1979). Although the representation of final consonants as appendices or extraprosodic constituents is not necessarily incompatible with the research explored here, these devices do not capture the probabilistic patterns of word-final consonants found in this position. For example, English has more voiceless obstruents than voiced obstruents in word-final position (Zamuner, 1998a, 1998b). By definition appendices or extraprosodic constituents are needed to account for why word-final positions are sometimes unrestricted in languages, as compared to word internal position. The research in this dissertation, however, finds that although languages might allow for an unrestricted set of word-final consonants, the distribution of these consonants is not random, and certain consonants are more frequent than others.

1.4.1.3.
Restrictions on analyses

By restricting analyses to word-final position in monosyllabic words, many potential confounds, such as those arising from positional markedness, are avoided. Positional markedness refers to markedness preferences that differ according to prosodic position. Because of this restriction, however, there are a number of other interesting factors that will not be addressed, which are worth investigating. Word-final consonants at the ends of unstressed syllables are restricted to a smaller set of consonants; moreover, they are restricted to a set of less marked consonants (Zamuner, 1997, 1998a, 1998b). This fits with theories of prominence (Beckman, 1998) in which prominent positions (stressed syllables) maintain contrasts and resists processes that apply in non-prominent positions (unstressed syllables).

There are further considerations of coda consonants with respect to the assignment of stress across languages and moraic structure (see Gordon, 2000). Zec (1988) presents data from Danish, Lithuanian, and Kwakwala suggesting that in these languages, CVC syllables are heavy when closed by sonorants (CVS) and light when closed by obstruents (CVO). In these three languages, sonorants behave as though they are moraic and obstruents behave as though they are not. Stress systems such as these might further skew the distribution of word-final consonants across languages. (See Beckman, 1998; Morén, 1998 and Zoll, 1998; for a discussion of similar issues.) One also finds, however, that there are more sonorants at the end of unstressed syllables than at the end of stressed syllables. This finding reflects the large number of syllabic consonants in this position, e.g., in the words: "bantam," "lightning," "captain," "maple" and "player." The proportion of sonorants is contrary to what one might predict based on the fact that some languages assign weight only to sonorant codas. Because unstressed syllables are light, one might predict that unstressed syllables in English would have fewer sonorant codas. The opposite pattern is found. (Also see den Ouden & Bastiaanse, 2001 for a discussion of the interplay of segmental and syllabic markedness in Dutch aphasics.) By restricting the analyses to just word-final consonants in stressed monosyllabic CVC words, confounds arising from these issues are avoided.

By excluding polysyllabic words, inventories of coda consonants in word-final position relative to word-medial position are also not considered. Research by Piggott (1999) has argued that there is a distinction between languages whose word-final consonants are syllabified as codas, and languages in which they are syllabified as onsets of empty-headed syllables. Moreover, Botma and van der Torre (2000) argue that in Dutch, sonorants are syllabified as codas, whereas obstruents are syllabified as onsets of empty-headed syllables. These possibilities are not address here.

There are many factors that affect the distribution and behaviour of coda consonants in languages. For example, there are interactions between different markedness preferences in coda position that are not compatible with each other. Languages prefer codas that are sonorant, that have coronal place of articulation, and that are voiceless, suggesting /l̥/ and /ɾ̥/ as common; however, the unmarked voice feature for sonorants is voiced. Obviously, given the many factors to consider and their possible interactions, it would be an extremely difficult task to examine word-final consonants with respect to a language's system as a whole. (See Coberly, 1985 for an examination of cross-linguistic and acquisition patterns of initial and final consonants.) Consider English, where it is not clear which information is relevant either for determining markedness or for determining the representation of final consonants. In an examination of the English lexicon, Zamuner (1998a) found that there are 21 codas permitted at the ends of CVC words, however, this set is reduced to 19 at the end of trochaic words. Furthermore, when one considers grammatical category, trochaic nouns permit a larger set of final consonants than trochaic verbs, and iambic verbs permit a larger set of final consonants than iambic nouns. These sorts of issues are set aside, and the starting point taken here is to provide an examination of word-final consonants at the end of CVC words.

1.4.1.4.
Place of articulation

Place of articulation refers to a system of classification based on where along the vocal tract the air stream is obstructed in the production of segments (Ladefoged, 1993). The three primary places of articulation are: labial (articulated with the lips), coronal (articulated with the tip or blade of the tongue), and dorsal (articulated with the back of the tongue). Each category contains finer distinctions; for example, coronal place of articulation contains both alveolar and alveolar-palatal articulations.

Evidence to support the hypothesis that there is a hierarchical organisation of place of articulation comes from cross-linguistic alternations that suggest there are classes based on place of articulation, and that these broad classes or features function as a unit in phonological processes. For example, in Ponapean, coronals are targets for place assimilation (13a), whereas labials and dorsals trigger assimilation (13b-c). Codas will be examined with respect to their place of articulation as labial, coronal, or dorsal.

(13)		Ponapean (Rehg & Sohl, 1981, cited in Rice, 1993).	
a.	nan-sed	nansed	*ocean*
b.	nan-par	nampar	*trade wind season*
c.	nan-kep	naŋkep	*inlet*

1.4.1.5.
Sonority

Acoustically, a segment's sonority refers to "its loudness relative to that of other sounds with the same length, stress, and pitch" (Ladefoged, 1993, p. 245). In articulatory terms, it refers to the approximate degree of aperture during a sound's articulation. Sonority also helps account for how sound sequences are organised into syllables. For example, in English, sonority is used to account for possible consonant sequences in word initial position, where "pla" is permitted and "*lpa" is not. To account for these data, a scale of sonority has been hypothesised, which governs how consonants and vowels are organised into syllables (Foley, 1972; Jespersen, 1904; Sievers, 1881 and others) (14):

(14) Sonority Hierarchy
less sonorous → more sonorous
Stops<Fricatives<Nasals<Liquids<Glides<Vowels

Languages then combine sounds based on their relative sonority. Thus, the case of "pla" is permitted because the sonority of the consonants rises towards the nucleus; whereas, "*lpa" is not permitted because the sequence has sonority peaks "l" and "a." (This system of government is known as the Sonority Sequencing Principle and the Sonority Sequencing Generalization, see Selkirk, 1984; Steriade, 1982 and others.)

This Sonority Hierarchy is argue to be innate (provided by UG), but how individual languages utilise this scale varies. A broad distinction is also made in the literature between sonorants and non-sonorants (obstruents: stops and fricatives). This approach is taken here; however, when relevant, I will describe the sonority of codas along the sonority hierarchy in finer detail.

1.5. Goal and outline of dissertation

Returning to the goal of the dissertation, I explore the place and sonority of word-final codas in the domains of cross-linguistic markedness, the distributions of codas in English and in coda acquisition in English. The aim is to determine whether children's productions reflect UG or the ambient language. The first possibility is that children's productions reflect innate UG (as evidenced in typology). Alternatively, children's productions could reflect the distributions in the input, which in turn reflect cross-linguistic markedness (although at times the specific language patterns might differ from cross-linguistic markedness). These two answers are not mutually exclusive; however, the latter has not previously been considered, and there is no comprehensive data analysis on the distributions of word-final codas or coda consonant acquisition in English.

Crucially, to determine whether children's productions reflect UG or whether they reflect properties in the input that also reflect UG, the analysis requires three types of data (Table 2).

Table 2 Data collected in dissertation and required for analysis

X-Linguistic Patterns	Language Specific Patterns: English	Child Language Patterns: English
(Chpt. 2)	(Chpt. 3)	(Chpts. 4–6)

The first type of data required is a characterisation of cross-linguistic markedness facts in word-final coda position. From here, one can determine the universal or unmarked preferences for place of articulation and sonority for codas. In Chapter 2, I present cross-linguistic research on codas, based on previously published research and on my own corpora studies. These data allow for the characterisation of the "universally" preferred coda consonants. The second type of data needed is the distribution of codas in English. Thus, Chapter 3 discusses the language-specific input for children exposed to English based on data from a variety of sources. These data allow us to determine the language-specific input for children acquiring English.

Lastly, data are needed from children acquiring English to determine whether children's coda productions pattern with UG preferences, or whether children produce the most frequent codas (Chapters 4–5). Chapter 4 comprises an analysis of coda consonant acquisition in English based on previous studies and data from CHILDES. In Chapter 5, I present an experiment designed to provide heretofore missing data on children's production of a variety of codas. The remainder of Chapter 5 presents evaluations of the UGH and the SLGH with respect to child language data.

In Chapter 6, I detail an experiment looking at a different type of frequency information, phonotactic probability. This experiment found that children were more likely to produce the same coda in high phonotactic probability non-words than in low phonotactic probability non-words. Both Chapter 5 and Chapter 6 provide evidence that children's production of coda consonants is best characterised by an input-based account of acquisition, where the distribution of sound patterns in the ambient language serve as the basis for phonological acquisition. A further discussion of these issues and a conclusion are given in Chapter 7.

CHAPTER 2
Cross-linguistic codas

2.1.
Introduction

In this chapter I explore cross-linguistic preferences for word-final codas in CVC words as a means of determining the universally unmarked features in this position and to establish the predictions of the Universal Grammar Hypothesis (UGH) for children's acquisition of coda consonants. To begin, I discuss the relevant data for determining cross-linguistic markedness in coda position. I then turn to an examination of previously published research and my own cross-linguistic statistical analyses of place of articulation and sonority in coda position.

2.2.
Data

The cross-linguistic data considered here come from previously published research and my own data collection of lexicons of CVC words from 35 languages (see Appendices I and II). Both types of data allow for an analysis of cross-linguistic word-final coda preferences. To obtain the CVC data, grammars and dictionaries were scanned for all or a randomised subset of CVC words. Loan words or words containing phonemes that appear only in loan words (when described by the grammar or dictionary as such) were excluded from the analysis. For example, in Telefol (Healey, 1977; Healey & Healey, 1977), words with final /s/ are loan words, and thus, any CVC word with a final /s/ was not included.

The set of possible codas and number of word-final codas for each language was then tabulated with respect to place and sonority. The size of the lexicons ranged between 33 to 1153 words (M=242.8, median=148). The size of the lexicons varied depending on the completeness of the dictionaries or grammars and depending on the number of CVC words within individual languages. For example, although the size of the CVC lexicon for Latvian contained only 64 words, three references were used to obtain these data: two grammars (Fennell & Gelsen, 1980; Nau, 1998) and one dictionary with approximately 37, 638 words (738 pages with approximately 51 words per page) (Turkina, 1982). It is assumed that the proportions of codas that exist in the samples are representative of the proportions of codas that exist in the individual languages as a whole. The 35 languages were chosen in order to provide data on CVC words from a variety of languages and sub-families. So as not to skew the results in favour of one family type, languages were restricted to one example from each sub-family. It is assumed that the sample is random. Many languages do not have coda consonants (such as many Native American Languages). In addition, although many languages permit codas, a large number of these languages have a disyllable minimal word requirement. Because the study was restricted to word-final codas in CVC words, these languages did not fit the criterion. The restricted distribution of languages with codas in CVC words and potential implications of this restriction, such as dependencies based on language families, is an interesting research question; however, it will not be explored here. (See Perkins, 1989 for a discussion on sample size and efficient sampling techniques in typological research.)

2.2.1.
Determining markedness

Recall that several types of data are used as evidence for cross-linguistic markedness: restricted universals, language processes, language change, patterns seen in child language and aphasia, and the distribution or frequency of sounds in languages. The research here is unique in that it determines the statistical patterns of word-final codas in CVC words.

The motivation for looking at statistical preferences for codas across languages begins with the observation that cross-linguistic markedness is reflected in the text frequency of languages (see also discussion in §1.2.1). Unmarked segments are generally more frequent than marked segments (e.g., Ferguson, 1963; Greenberg, 1966, Trubetzkoy, 1931, 1939/1969; Zipf, 1935/1965). For example, voiceless obstruents are less marked than voiced obstruents; thus, voiceless stops should have a higher frequency than voiced stops in any one language. This is true for at least Beijing Chinese, Cantonese Chinese, Danish, Burmese, Czech, Dutch, Finnish, Italian, English, Hungarian, Bulgarian, Russian, Spanish, Greek, Latin and Sanskrit (Zipf, 1935/1965). The assumption is that a cross-linguistic statistical analysis of word-final codas will allow for a clarification of what is unmarked and marked in this position. Because cross-linguistic markedness patterns are interpreted here as evidence for UG, an analysis of these patterns will subsequently allow for a prediction of what patterns are expected to emerge in children's productions of coda consonants. I now turn to a discussion of how statistical patterns will be established, by considering the data on which frequency is calculated and considering what frequency analyses are appropriate.

2.2.2.
Types of data

There are different types of data on which to calculate frequencies, e.g., based on written or spoken corpora and/or lexicons. For example, the Prague School of Linguistics emphasised the pertinence of looking at the distribution of a language's sounds not only in terms of absolute numbers of occurrences, but also in terms of the "ratio of these figures to the figures of occurrences theoretically expected on the basis of combinatory rules" (Trubetzkoy, 1939/1969, p.257). For an early discussion of these issues, also see Krámský 1946–48, 1959a, 1959b, cited in Greenberg 1974, p.47, and Trubetzkoy (1939/1969, chap. VII).

It is also possible to examine type versus token frequencies. Type word counts refer to the number of unique words in a corpus or a dictionary. Token word counts refer to the total number of words. For example, in the following sentence, there are 20 word types ("to," "that," and "is" are repeated) and 23 word tokens:

(1) "Of course, an alternative to both hypotheses, and one that is often not considered, is that language was given to us by aliens." (Reverend Ethan Cox)

The data presented here are based on type word counts of lexicons from 35 languages. Ideally, token word counts from written or spoken corpora should also be examined. Unfortunately, these data are not currently available from a wide enough range of languages.

2.3.
Types of analyses

Two different frequency analyses were done on the data. The first analysis will be referred to as an Expected Frequency Analysis (EFA) and the second will be referred to as an Actual Frequency Analysis (AFA).

2.3.1.
Expected Frequency Analyses (EFA)

This analysis determines whether the words of a language contain a specific phonological element more than expected by chance. In this case, chance is based on the proportion of codas that have a specific phonological feature to the total inventory of codas, multiplied by the number of words in the lexicon. In other words, the number of codas with a specific phonological feature (N_{PFC}) are divided by the number of possible codas types in the language (N_C), then multiplied by the number of words in the lexicon (N_L) (2). The result gives the number of words one would expect to find in the lexicon with a specific phonological feature. The result of the EFA is then compared to the number of words in the language that do, in fact, have the coda type in question. By comparing expected versus observed frequencies, the EFA takes into consideration the fact that there are more codas with certain features in the inventory than other places of articulation.

(2) Calculation for Expected Frequency Analysis

$$EFA = \frac{N_{PFC}}{N_C} * N_L$$

To illustrate take Dutch (Baayen, Piepenbrock & Rijn 1993), which has eleven codas in the inventory (p, f, m, t, s, n, r, l, k, x, ŋ); five of these bear the feature coronal. By chance, one would expect a proportion of . 45 coronal codas (5 coronal codas divided by 11 codas in the inventory) in the total lexicon (1153 words). Thus, it is expected that 524 words should end in coronal codas (.45 by 1153), as shown below:

(3) Dutch Expected Frequency Analysis

$$\frac{N_{PFC}}{N_C} * N_L$$

$$= \frac{5}{11} * 1153$$

$$= .45 * 1153$$

$$= 524$$

In this example, however, 669 words end in coronal codas (see Table 1). In general, by considering expected frequencies, this controls for the fact that languages have more coronal codas in their inventories than codas with other places of articulation.

Table 1 Expected Frequency Analysis of coronal coda place in Dutch

Expected frequency	Observed frequency
524	669

An EFA was computed for each language and data from the 35 languages were combined to determine the overall cross-linguistic effect. Chi-square goodness-of-fit tests determined whether across the 35 languages there were significantly more words with codas of a specific phonological feature than expected, as compared to the observed number of words in the languages that do have the coda type in question.

2.3.2. Actual Frequency Analyses (AFA)

The second way the cross-linguistic data were analysed was to determine whether there was a greater number of codas with a certain phonological feature across the 35 languages. This analysis differs from the EFA in that it establishes whether the number of words containing codas with a specific phonological feature is greater than the words containing different features. The AFA does not take into account whether languages have more codas with a specific phonological feature in proportion to other codas in the language. In this sense, the AFA simply measures which are the most frequent codas in a language. To illustrate this type of analysis, consider again place of articulation in Dutch: 669 of the lexicon's words end in coronals, whereas 484 of the words end in labials and dorsals combined (Table 2).

Table 2 Actual Frequency Analysis of coronal coda place in Dutch

Coronal POA	Other POA
669	484

An AFA was performed on each language. To determine whether there were significantly more codas with a specific phonological feature than other features. A t-test was performed, which collapsed data across the 35 languages to provide the general cross-linguistic pattern. It would also be possible to separate labial and dorsal codas from coronal codas, however, combining them provides a more generous and less biased estimate. The reason for this is that there are many coronal codas in languages' inventories, but comparatively few labial and dorsal codas.

To summarise, both EFAs and AFAs were performed on the cross-linguistic data to determine preferences for place and sonority in word-final coda position. The EFA determines whether languages have more codas with a certain feature than expected by chance based on the inventory of codas in the language. In the case of coronal place of articulation, this calculation takes into consideration the fact that there are more coronal codas in most languages' inventories than other places of articulation. The AFA determines whether the number of words with a specific phonological feature is greater than words with different features. This analysis does not consider the unequal distributions of features in codas, as seen in the case of coronal where there are more coronal codas in languages' inventories than other codas.

2.4. Place of articulation in coda position

In this section, I review previous research on place of articulation in coda position, and present my own cross-linguistic frequency analyses of place of articulation in word-final position.

2.4.1. Previous research on place of articulation in coda position

Regardless of syllable position, coronal consonants are unmarked (see Paradis & Prunet, 1991), and cross-linguistically, coronals are the most frequent (Maddieson, 1984). For example, the coronal nasal /n/ is much more common than the labial nasal /m/, and some languages have *only* /n/ (e.g., Arapaho (8), Chipewyan, Tlingit, Wichita, Yuchi, but cf., Taoripi which has only /m/). Liquids are also most frequently coronal, and some languages have only coronal liquids (e.g., Fuchow, Koiari, Mixtec, and Sedang, Maddieson, 1984). In addition, some languages restrict word-final codas to coronal place, as in Finnish (Yip, 1991).

Place in coda position also tends to be restricted to geminates or homorganic nasals that share place with the following onset (Coda Condition; Itô, 1986). For example, in Lardil (4), labials and velars appear in coda position only if they share place with a following onset (a and b), but as seen in (c–f), no such restriction applies to coronals (except with lamino-dentals; from Rice, 1993).

(4)		Lardil (Wilkinson, 1988)
a.	ŋam.pit	*humpy*
b.	kuŋ.ka	*groin*
c.	kar.mu	*bone*
d.	pir.ŋa	*woman*
e.	kan.tu	*blood*
f.	rel.ka	*head*

The argument is that coronal is the unmarked place of articulation regardless of syllable position; moreover, there are positional constraints on segments such that coronals tend to be the unmarked place segment in coda position, whereas labials and dorsals are restricted in this position. Consequently, one should find a preference for coronal codas across languages. (I will not consider here whether coronals are underspecified for place; for a discussion see Lombardi, 1996.)

In some languages velar is the unrestricted place feature in coda position (Rice, 1993). To determine whether a languages' unmarked place feature is coronal or dorsal requires an analysis of a language's phonotactics and phonological processes, which is beyond the scope of the dissertation. One can argue, however, that although both coronal and dorsal are unmarked in coda position, coronal is the least marked of the two because it is most frequently the unmarked place of articulation in coda position. For a similar argument, see Fonte (1996) who claims that although there is no universally unmarked place of articulation, evidence from acquisition suggests that coronal is less marked than dorsal or labial (Goad, 1997b). Further evidence from a variety of languages beyond Dutch and English are needed, however, to determine the general acquisition pattern.

2.4.2.
Place of articulation in coda position: EFA

Recall that an EFA determines whether languages have significantly more codas with a specific phonological feature than expected by chance, taking into account the inventory of codas in a particular language and the number of words in the lexicon. Data from the 35 languages were combined. Here a chi-square goodness-of-fit test revealed that there were more coronal codas than expected, χ^2 (34, N=7906) =209.29, $p<.001$. To determine whether the pattern of results was significant for the individual languages and to determine whether the overall cross-linguistic pattern held for each language, a chi-square goodness-of-fit test was run on the data from each language. Results indicated that 21 of the 35 languages had significantly more coronal codas than expected, and four languages had significantly fewer coronal codas than expected. (See Appendix III for a summary of how each language patterned.) As a whole, coronal emerges as the cross-linguistically preferred place of articulation in coda position.

2.4.3.
Place of articulation in coda position: AFA

The second way the cross-linguistic data were analysed was to determine whether there was a greater number of coronal codas than other places of articulation across the 35 languages. This analysis involved comparing the number of words in each lexicon that ended in coronal codas to the words ending in other places of articulation, and then combining the results. Across the 35 languages, there were more coronal codas (M=133.8, SD=11.67) than non-coronal codas (M=92.03, SD=11.26). This difference was significant, $t(34)$ 2.67, $p<.01$, two-tailed. Twenty-seven out of the 35 individual languages had more words ending in coronal codas than other places of articulation. (See Appendix IV for summary of how each language patterned.)

2.4.4.
Summary of place of articulation in coda position

Data from previous research and from frequency analyses support the claim that there is a preference for coronal codas cross-linguistically (Table 3).

Table 3 X-Linguistic preference for coronal codas

Previous X-Linguistic Research	Expected Frequency Analysis	Actual Frequency Analysis
Coronal	Coronal	Coronal

It is important to note, however, that as seen with the individual languages in the EFA analyses, some languages do not prefer coronal codas. Thus, coronal place of articulation is not an unrestricted universal, but rather, a statistical universal, in which the generalisation describes languages' tendencies to pattern a particular way (see §1.2.1 for a discussion of different universals). Despite languages that pattern in the opposite direction, it is reasonable to assert that coronal is the unmarked place of articulation in coda position given that coronals are significantly more likely to be unmarked than not.

2.5.
Sonority in coda position

The other parameter to be discusses in this chapter is sonority. As with place of articulation, I present previous research on universal preferences for sonority in coda position, and then provide EFAs and AFAs for cross-linguistic data.

2.5.1.
Previous research on sonority in coda position

Across languages, the preferred syllable shape has a sonorous ending; thus, the preferred syllable ends with a final vowel. Yet, when final consonants are present, the preferred syllable-final consonants are sonorant consonants, such as nasals and liquids:

> The simplest syllable is one with the maximal and most evenly distributed rise in sonority at the beginning and the minimal drop in sonority (in the limit case, none at all) at the end (Clements, 1990, p. 303).

This preference is found in typological surveys where languages tend to favour sonorant codas over obstruent codas. For example, in Manam the only possible codas are nasals (5). In Italian (6) and Japanese the only non-geminate codas are restricted to sonorants (see Itô, 1986), and in Beijing Chinese the only possible codas are /n ŋ ɹ/ (Blevins, 1995).

(5)	Manam (Lichtenberk, 1983)
em.be.ʔɪ	*sacred flute*
maŋ	*bird*
ma.daŋ	*Madang*
lun.ta	*moss*
ma.la.boŋ	*flying fox*
raŋ.gu.ma	*who, (pl)*

(6)	Italian
sɛl.və	*wood*
tom.bi.no	*manhole cover*
ɛr.to	*steep*
kon.te.ə	*county*
ær.rɛ.sto	*arrest*
græp.pə	*cramp, grappa*

Because of data like those in (5) and (6), sonorant codas are argued to be favoured over non-sonorant codas in languages. Additionally, there is argued to be an unrestricted universal in language, such that the presence of an obstruent coda implies the presence of a sonorant coda, repeated in (7) (Clements, 1990; Eckman & Iverson 1994; Fonte, 1996; den Ouden & Bastiaanse, 2001).

(7) Unrestricted universal

X → Y

obstruent codas		sonorant codas	
Type 1	+	+	occurs
Type 2	−	−	occurs
Type 3	−	+	occurs
Type 4	+	−	*does not occur*

The pattern previously observed in (7) is supported by an examination of the 35 languages where all languages with obstruent codas had sonorant codas as well. Also with the languages that restricted codas to a certain type, the restrictions favoured sonorants over obstruents. Among the 35 languages examined, two of these allowed only sonorants in coda position: Salt-Yui and Mbay. Notice that the observation in (7) is based on previous data and analyses in which medial and final codas have been collapsed. There is an absence of data on codas solely in word-final position. Because the distribution of final codas differs from those in word-internal position (Harris, 1994; Piggott, in press; Itô, 1986), it was necessary to determine the sonority preferences of word-final codas. Statistical analyses of lexicons from 35 languages provided these missing data.

2.5.2. Sonority in coda position: EFA

The first test for sonority preferences in codas was to determine whether languages show a preference for sonorant codas (as suggested above). An EFA was performed in which the observed number of sonorant codas was compared to the expected number of sonorant codas across the 35 languages considered. A chi-square goodness-of-fit test revealed that there were more sonorant codas than expected by chance, χ^2 (34, N=7906)=197.78, p<.001. Furthermore, a chi-square goodness-of-fit test was run on the data from each language to determine whether the pattern of results was significant for each language, and to determine whether the overall cross-linguistic patterns held for each individual language. Results indicated that 13 languages had significantly more sonorant codas and 4 languages had significantly fewer sonorant codas than expected. (See Appendix V for a summary on how each language patterned.) Importantly, the overall pattern across languages was that there were significantly more sonorant codas than predicted. Because one also finds languages in which there are fewer sonorant codas than expected, this raises the possibility that languages exist in which marked structures are more frequent than unmarked structures. Not all languages show a preference for sonorant codas; thus, the statistical analyses illustrate that markedness is not absolute with respect to sonority preferences in coda position, and in this sense, the preference for sonorant codas is a statistical universal. (See §1.2.1 for a discussion of various universal types.)

Some arguments have been made for the special status of nasals in coda position (Fonte, 1996; Piggott, in press). Thus, it is possible that the sonority preference found across the 35 languages was carried by languages' preferences for nasal codas, rather than sonorant codas. An examination of the sonority preference across the 35 languages revealed although 15 languages had more nasal codas than expected (Arabic, Armenian, Cambodian, Cantonese, Daga, Dehong, Dehu, Doyayo, Hebrew, Jacaltec, Rhade, Thai, Tibetan, Totonac, and Ukranian), there were 20 languages with more liquid codas than expected (Alawa, Brahui, Creole, Daur, Dehu, Doyayo, Dutch, Hebrew, Jacaltec, Kurdish, Lisu, Mam, Mbay, Munari, Nivkh, Pima Bajo, Pulaar, Salt-Yui, Telefol, and Turkmen). Thus, the special status of nasals in coda position is not supported by these data and/or by this type of analysis.

2.5.3. Sonority in coda position: AFA

The cross-linguistic data were also analysed using an AFA to determine whether there were significantly more sonorant than obstruent codas in the languages' lexicons. Unlike with the EFA, no preference was found for languages to have either sonorants or obstruents. There were more obstruent codas (M=132.8, SD=12.03) than sonorant codas (M=111.06, SD=11.46); however, this difference was not significant, $t(34) = 1.09$, n.s., two-tailed.[3] Twenty-two languages had more words ending in obstruent codas than words ending in sonorant codas. See below for a discussion of this apparent anomaly, and see Appendix VI for a list of the languages and how they patterned.

2.5.4. Summary of sonority in coda position

Cross-linguistically the preferred syllable coda is one of high sonority, based on previously published research and on the EFA (Table 4).

Table 4 X-Linguistic preference for sonorant codas

Previous X-Linguistic Research	Expected Frequency Analysis	Actual Frequency Analysis
Sonorant	Sonorant	no preference

The evidence from previous research found that sonorants are universally unmarked in coda position, such that the presence of obstruents entails the presence of sonorants in coda position. In addition, an EFA of sonority in coda position revealed that there are more sonorants than expected. This conclusion was based on an analysis that took the number of words that were expected to have sonorant codas compared to the actual number of words ending in sonorant codas. Finally, although the AFA did not reveal a significant difference between the number of sonorant versus obstruents in languages, this is not surprising given the nature of languages' inventories. Languages' phoneme inventories are largely composed of obstruents (voiced and voiceless stops and fricatives with three different places of articulation), whereas sonorants are restricted to nasals (usually just voiced nasals with three different places of articulation) and liquids (which are almost always restricted to coronal place of articulation). Because languages have a larger set of obstruent phonemes than sonorant phonemes, this accounts for why a greater number of obstruents than sonorants are found across languages, despite sonorants being universally preferred in this position.

2.6. Universal Grammar Hypothesis for coda acquisition

This chapter determined the cross-linguistic preferences for coda consonants with respect to their place of articulation and sonority. Both data from previously published research and data from lexicons of CVC words from 35 languages were examined. The results are that coronal place of articulation is preferred in coda position and languages favour sonorant codas over obstruents. Based on these results, the conclusion

[3] Results from Coberly (1985)'s examination of the inventory of initial and final consonants across 58 languages found that obstruents are favoured in initial position over final position, which is consistent with Clements, 1990.

drawn is that the unmarked codas have coronal place of articulation and are sonorant. Given the cross-linguistic data, one is now in the position to operationalise the predictions of the UGH for English-learning children's early coda productions. The English coronal and sonorant codas are taken to characterise this preference (Table 5), as this will allow for a direct comparison between the distribution of codas in English and child language data from English, which are established in the following chapters.

Table 5 X-Linguistic patterns established in Chapter 2

	X-Linguistic patterns
CORONAL:	t, d, s, z, ʃ, ʒ, tʃ, dʒ θ, ð, n, l, r
SONORANT:	m, n, ŋ, l, r

The cross-linguistic patterns established in this chapter are taken to reflect UG. Thus, if language acquisition is governed by innate properties of UG, the preference for coronal and sonorant codas should be mirrored in early child language. The UG hypothesis for coda acquisition is as follows:

(8) Universal Grammar Hypothesis (UGH_1): Children will produce coronal codas (t, d, s, z, ʃ, ʒ, tʃ, dʒ θ, ð, n, l, r) more than labial or dorsal codas, and sonorant codas (m, n, ŋ, l, r) more than obstruent codas in word-final position.

The predictions given in (8) are taken as a starting point for a UG hypothesis of coda consonant acquisition. However, it would be unlikely to find the affricates /tʃ/ and /dʒ/ (for example) among children's initial coda productions, despite the fact that they are coronals. These segments require sophisticated articulation skills and are therefore marked in terms of manner. This raises the question of how to fully determine the markedness of any particular segment given that some properties of a segment may be unmarked whereas other properties of the same segment may be marked. Because of these complicating factors, it is not clear how an exact prediction of coda consonant acquisition should be formulated. To formulate UG predictions, one needs to consider all the possible cross-linguistic characteristics of codas and how they interact. Such an enterprise is beyond the scope of this dissertation. Despite these problems with operationalising a UG hypothesis, a reasonable starting point for the UG based hypotheses of coda acquisition (based on previous implementations of UG in acquisition) is that the UGH predicts children's initial coda productions should bear coronal place of articulation over labial or dorsal place of articulation and will preferably be sonorant over obstruent. Still the UGH can be revised to exclude those sounds known to be late acquired or marked along other dimensions (9).

(9) Universal Grammar Hypothesis (UGH_2): Children will produce coronal codas (t, d, s, z, n, 1, r) more than labial or dorsal codas, and sonorant codas (m, n, ŋ, 1, r) more than obstruent codas in word-final position.

Although it would also be possible to collapse these hypotheses and argue that the unmarked coda is a coronal sonorant, e.g., /n/, one must also consider the unmarked voice feature in coda position, which is voiceless. Combining the three preferences, the prediction would be that the unmarked coda would be a voiceless coronal sonorant, e.g., /n̥/ However, voiceless sonorants are cross-linguistically marked (Maddieson, 1984). Thus, the unmarked features in coda position are not compatible. Although voice is not considered in these analyses, the hypotheses have not been combined because this would give an inaccurate characterisation of the cross-linguistic preferences in this position. See §1.4.1.3 for a further discussion. The next step is to examine English frequency data, which establishes patterns of coda consonants in the input for children acquiring English.

CHAPTER 3
English codas

3.1.
Introduction

This chapter presents an analysis of the distribution of English coda consonants from a number of sources with the goal of establishing the phonological input for children acquiring English. To this end, English word-final codas in CVC words were examined. The results allow for a characterisation of the Specific Language Grammar Hypothesis (SLGH) for children acquiring English. The details of the chapter are as follows. First, previous studies of phoneme distributions in English are reviewed. I then describe the English sources on which the present coda analyses are based. The SLGH is then defined for children's acquisition of English coda consonants. I conclude with a discussion of the units of analyses used in the different theories of acquisition contrasted in this dissertation.

3.1.1.
Previous studies of phoneme distributions in English

There are a number of studies in the literature that have looked at the distribution of phonemes in English (e.g., Dewey, 1923; Roberts, 1965); however, no comprehensive study has examined the distribution of codas with respect to markedness, nor have these investigators necessarily separated codas from onsets in their analyses. Researchers that do make this latter distinction have focused on comparing onset and coda clusters, while others have explored the predictability of onsets and codas from the syllable's vowel. In the analysis presented here, I neither compare coda inventories to onset inventories, nor consider the effects of a syllable's vowel in distributions of word-final consonants. For these types of studies see Coberly, 1985; Coleman and Pierrehumbert, 1997; Kessler and Treiman, 1997, and Pierrehumbert, 1994.

Some studies that have distinguished between onset and coda position fail to control for other prosodic positional effects. Thus, the inventory or frequency of codas is not distinguished according to words' syllable length and syllable stress. For example, a study by Mines, Hanson and Shoup (1978) calculated the frequency of phonemes in natural spoken English in word initial, word-medial and word-final positions; however, word-final consonants were not further divided into those that occur at the end of single versus multiple syllable words, or into word-final consonants that occurred in unstressed or stressed positions (e.g., at the end of trochaic and iambic words, respectively). A study by Dobrich and Scarborough (1992) examining the distribution of phonemes in children's early words has similar limitations. The factors neglected in these studies are important given that the inventories of final consonants differ according to prosodic position (Zamuner, 1998a, 1998b). For these reasons, previous analyses of coda frequencies in

English were not appropriate, and it was necessary to establish the distribution of word-final codas in English CVC words.

3.2.
English data

Data were taken from four sources: the Random House Dictionary, the new Merriam-Webster pocket dictionary, the MacArthur Communicative Development Inventory (CDI), and a corpus of Child Directed Speech (CDSC). Different databases were used to determine whether the patterns of coda consonants in English would vary across databases. Analyses based on a variety of sources allow for a wider interpretation of the results, and allow for a determination of whether the same patterns are found in adult lexicons and child directed speech. CVC words from the lexicons or corpus were compiled and the number of word-final consonants for each of the English phonemes was recorded. Where possible, the number of word-final codas based on token and type word counts was also calculated. Each data source will now be discussed in detail.

3.2.1.
Random House Dictionary

Data from the Random House Dictionary (Flexner, 1987) were taken from a study by Kessler and Treiman (1997), which looked at the syllable structure and distribution of phonemes and features of uninflected CVC English words. Kessler and Treiman used a broad definition of inflection, such that they also excluded any words containing cranberry morphemes; for example, the words “this” and “then” were excluded because “th” could be analysed as a demonstrative morpheme. Although children do not receive an input that is composed of uninflected words, Kessler and Treiman investigated this restricted set with the intention to further extend the analysis to both uninflected and inflected words and other word positions. They excluded non-American English words, “Thus we omitted words with foreign phonemes or accented letters, foreign measures, and places and ethnic names that were not obviously Anglicized” (Kessler & Treiman, 1997, p. 299). They also treated /ɚ/ as a single phoneme; thus, words like “fern” constituted a CVC word. Kessler and Treiman’s data are based on type word counts from the Random House Dictionary.

3.2.2.
Webster’s Dictionary

Data from the Webster’s Dictionary were obtained from an electronic copy of the new Merriam-Webster pocket dictionary (1964). Phonetic transcriptions in the electronic version were obtained from other sources. Following Kessler and Treiman, /ɚ/ was also treated as a single phoneme. However, in contrast to their analyses, no words were excluded from the final list of CVC words. Type word counts for this lexicon were examined.

3.2.3.
MacArthur Communicative Development Inventory (CDI)

The MacArthur Communicative Development Inventory (CDI) (Fenson, et al., 1993), is a parent survey that lists approximately 475 typical words in young children’s vocabularies. This “lexicon” allowed for an analysis of the distribution of codas in a subset of English words that are familiar to children. To obtain these data, all CVC words were taken from the list. Data on coda consonant frequencies are taken from type word counts.

3.2.4.
Child Directed Speech Corpus (CDSC)

The last database was a corpus of child directed speech (CDSC). To create this corpus, parental speech was taken from CHILDES studies involving children between the ages of 19 to 28 months; the studies are listed in Table 1.

Table 1 Studies from which data were obtained for the CDSC

Corpus	Reference
Bates Corpus	(Bates, Bretherton & Snyder, 1988; Carlson-Luden, 1979)
Clark Corpus	(Clark, 1979)
Demetras and Umbreit Corpus: Trevor	(Demetras, 1989)
Demetras Corpus: Working Parents	(Demetras, 1989a, 1989b)
Higginson Corpus	(Higginson, 1985)
MacWhinney Corpus	(MacWhinney, 1999, 2000)
Warren-Leubecker Corpus	(Warren-Leubecker, 1982; Warren-Leubecker & Bohannon, 1984)

Parental speech was obtained only from studies conducted in naturalistic settings. This ensured that there would not be an unnaturally high frequency of sounds from specific words. For example, some studies in the Bernstein-Ratner Corpus use a pre-determined set of toys in all sessions. The resulting CDSC consisted of 143, 496 word tokens, with 4,891 unique types. Phonetic transcription was obtained from the electronic version of the Webster's Dictionary of American English, described in §3.2.2. In some instances the dictionary did not contain the phonetic transcription for a word in the CDSC, e.g., for the plural word "boys." All these remaining cases were transcribed by hand, based on other words in the dictionary. Parental speech was all speech excluding the productions from the child under investigation; therefore, "parental speech" includes both adult and child speech (from siblings) and adult-directed and child-directed speech.

All proper names were then excluded. It was assumed that this type of input varies greatly across children due to children having different names from popular culture, e.g., "Tom" and "Hoth" (the former is a subject's name, and the latter refers to the movie Star Wars and the ice planet on which the rebel base is located at the beginning of Empire). The resulting CVC word corpus consisted of 40,822 tokens, with 604 unique types.

3.2.5.
Summary of sources

To recap, data were taken from two official English lexicons (the Random House Dictionary and the new Merriam-Webster pocket dictionary), from a "lexicon" of typical words familiar to young children (CDI), and from a CDSC compiled from CHILDES.

Table 2 Rank order of English codas from different sources

	Random House	Webster	CDI	CDSC type	CDSC token
most	l	t	t	t	t
frequent	n	l	k	d	r
	t	n	d	n	n
	k	k	r	l	d
	d	d	s	k	z
	m	p	p, n	z	k
	r	m	l	p	s
	s	r	m	r	l
	p	s	g	m	m
	f	z	z, tʃ	s	v
	g	g	ʃ	θ	ʃ
	b	tʃ	v, θ, ŋ	ŋ	g
	tʃ	v	f	f, v	p
	z	b, f	b	g	θ
	v	θ	ʒ, dʒ, ð	ʃ, tʃ	ŋ
	ŋ	dʒ		b	tʃ
	ʃ	ʃ		dʒ	f
	dʒ	ŋ		ʒ, ð	dʒ
	θ	ð			b
least	ð	ʒ			ʒ, ð
frequent	ʒ				

3.3.
Distribution of codas in English

The rank order of English codas from each of these sources is given in Table 2, and the number and percentage of word-final codas in CVC words from each source is summarised in Table 3.

Consistency can be found in the rank order of codas given in Table 4. Across all databases, within the 10 most frequent codas are /t, d, l, n, k, m, r, s/, and within the 10 least frequent codas are /b, ŋ, dʒ, ð, ʒ/. There are then three primary considerations before developing the SLGH for coda consonant acquisition. Although there is consistency in the distribution of codas across the different databases, one first needs to determine which data are most appropriate for forming predictions on acquisition. One can choose between a lexicon and a spoken corpus, and frequencies can be calculated on type or token word counts. The predictions here will be based on the token counts of the CDSC for the following reasons: this source provides the closest approximation to children's input between the ages of 19 and 28 months, and token frequency has been shown to predict acquisition more accurately than type frequency (Marchman & Plunkett, 1989).

The second consideration is the type of analysis to perform on the English data. It is possible to determine whether English reflects the universal preference for coronal place of articulation and sonority in coda position, as seen with cross-linguistic data in Chapter 2. To this end, both EFAs and AFAs were performed on the English data from each source. Results of the EFAs indicated that English has more coronal codas than expected. AFAs indicated that there are more coronal codas than other places of articulation (See Appendices VII and VIII). If children learning English initially produce coronal codas, it would not be clear whether children's productions reflect the universal place of articulation preferences, or the distribution of

Table 3 Number and percentages (in parentheses) of word-final codas in English CVC words from different sources

Coda	Example	Random House	Webster's	CDI	CDSC type	CDSC token
p	chip	112 (5.6)	77 (5.93)	10 (6.58)	38 (6.29)	458 (1.12)
b	pub	62 (3.1)	41 (3.16)	1 (0.66)	6 (0.99)	51 (0.12)
f	rough	68 (3.4)	41 (3.16)	2 (1.32)	15 (2.48)	85 (0.21)
v	shove	54 (2.7)	42 (3.23)	3 (1.97)	15 (2.48)	1384 (3.39)
t	shot	204 (10.2)	143 (11)	23 (15.1)	82 (13.6)	10383 (25.4)
d	dude	142 (7.1)	97 (7.47)	14 (9.21)	61 (10.1)	3337 (8.17)
s	pass	116 (5.8)	69 (5.31)	12 (7.89)	31 (5.13)	2566 (6.29)
z	noise	58 (2.9)	56 (4.31)	5 (3.29)	42 (6.95)	3106 (7.61)
ʃ	bash	44 (2.2)	28 (2.16)	4 (2.63)	10 (1.66)	775 (1.9)
ʒ	beige	5 (0.25)	3 (0.23)	0	0	0
tʃ	coach	59 (2.95)	43 (3.31)	5 (3.29)	10 (1.66)	157 (0.38)
dʒ	dodge	41 (2.05)	29 (2.23)	0	4 (0.66)	57 (0.14)
θ	teeth	39 (1.95)	33 (2.54)	3 (1.97)	21 (3.48)	425 (1.04)
ð	writhe	14 (0.7)	11 (0.85)	0	0	0
k	puck	182 (9.1)	107 (8.24)	22 (14.5)	54 (8.94)	3055 (7.48)
g	league	67 (3.35)	47 (3.62)	6 (3.95)	14 (2.32)	593 (1.45)
m	game	127 (6.35)	74 (5.7)	7 (4.61)	32 (5.3)	2203 (5.4)
n	zone	207 (10.3)	127 (9.78)	10 (6.58)	58 (9.6)	4161
ŋ	wing	46 (2.3)	27 (2.08)	3 (1.97)	17 (2.81)	362 (0.89)
l	goal	230 (11.5)	134 (10.3)	9 (5.92)	57 (9.44)	2400 (5.88)
r	cheer	124 (6.2)	70 (5.39)	13 (8.55)	37 (6.13)	5264 (12.9)
Total		2001	1299	152	604	40822

place in the ambient language. Similarly, both EFAs and AFAs were done for sonority in English CVC words from all sources. The EFAs showed that English has more sonorant codas than expected. The AFAs found that English has more obstruent than sonorant codas (see Appendices IX and X). Interestingly, there are different coda preferences depending on the analysis; there is a larger proportion of sonorant codas than expected, but a significantly greater number of obstruent than sonorant codas in CVC words. As in Chapter 2, this latter finding can be accounted for by appealing to English's coda inventory, which has more obstruents (16) than sonorants (5). Examining children's production of sonorants and obstruents in English then provides an interesting test case for the issues explored in this dissertation. If children initially produce obstruents in coda position, this would suggest that children produce the most frequent codas in their language, rather than the unmarked codas.

The third concern is whether to base the predictions for children's acquisition of coda consonants on the raw frequency or on the frequency-based rank order of codas, both of which are possible. For ease of exposition, I will adopt a rank order approach. Predictions of children's acquisition of coda consonants in English based on the rank order of codas are given in Table 4 below.

Table 4 Language specific patterns for English

Language Specific Patterns: English
t>r>n>d>z>k>s>l>m>v>ʃ> g>p>θ>ŋ>tʃ>f>dʒ>b>ʒ, ð

3.4.
Specific Language Grammar Hypothesis for coda acquisition

The input-based hypothesis for children's acquisition of coda consonants based on the rank order of token coda counts of the CDSC is stated below.

(1) Specific Language Grammar Hypothesis (SLGH): Children will produce the most frequent codas in word-final position: t>r>n>d>z>k>s>l>m>v>ʃ>g>p>θ>ŋ>tʃ>f>dʒ> b>ʒ, ð.

The prediction of children's acquisition of coda consonants based on the SLGH is similar to the UGH with respect to place of articulation in coda position because some of the most frequent codas in English have coronal place. For example, both hypotheses predict that children will show an early preference for the codas: /t, n, r, d/. The SLGH differs from the UGH in that not all of the frequent codas in English have coronal place of articulation nor are they all sonorants, for example /k/ is a frequent coda, yet it is an obstruent with dorsal place of articulation.

3.5.
Comparing the UGH and the SLGH

Under the present characterisation of UG, the prediction is that the individual segments children initially produce will be coronal and sonorant. The interpretation of the input-based account of phonological acquisition offered here is that children will initially acquire the most frequent codas, based on the distribution of coda exemplars in the input. The UGH makes more general predictions (i.e., that the initial segment will be coronal/sonorant, but not a particular coronal/sonorant) than that of the SLGH (i.e., that the initial segments will be, specifically, t>r>n>d>z etc). The question then arises as to whether the more specific predictions of the SLGH can be abstracted to produce more general predictions so the two theories can be more directly compared. One clear difference between the two approaches, however, is that the SLGH predicts that frequency distributions will be random with respect to phonetic features like coronal and sonorant. In contrast, the UGH predicts that coronals as a group will be more frequent than non-coronals, and that sonorants as a group will be more frequent than obstruents. If adult input and child data both show a random distribution of consonants features, and if the rank orders of the two types of data are significantly correlated, this will be evidence for the SLGH. In contrast, if the adult input and child language data show groupings of coronal and sonorant codas, then children's coda productions will be consistent with both the UGH and the SLGH. To address this issue, the next section explores whether the distribution of English codas demonstrate feature-based grouping, i.e., whether the most frequent codas in English group together by a particular place of articulation or sonority.

3.5.1.
Place of articulation and the distribution of codas in English

This section explores whether there is a most frequent place of articulation in the rank order of English codas. To determine this, the rank order of English codas based on the token word counts from the CDSC were analysed for evidence of coronal, dorsal or labial place of articulation. A Kruskal-Wallis one-way analysis of variance (ranks) was done. This test takes the median rank of groups to determine whether they belong to a single population. To compute the Kruskal-Wallis statistic, coda consonants were ranked according to their frequency. Because there are 21 possible codas in English, ranks ranged from 1 to 21, where 21 was assigned to the most frequent coda /t/, and 1.5 assigned to the two least frequent codas /ʒ/ and /ð/. Table 5 provides the rank of English codas based on the token word counts of the CDSC. Table 6 categorises English codas based on place of articulation and also provides the sum and mean rank for codas with coronal, dorsal and labial place of articulation. This test measures whether the groups of coronal, dorsal and labial codas are from the same population. If so, then the average ranks should be similar. The null hypothesis is that place of articulation will not be reflected in the rank order of English codas, and results support the null hypothesis. The groups of coronal, dorsals and labial codas were not significantly represented in the rank order of English codas, H (2, N=21)=1.22, *p*=.54.[4] That is, despite the fact that the average rank of coronal codas in English is higher than that of dorsal or labial codas (12 versus 11 and 8.4, respectively, see Table 6), this difference is not statistically significant.

3.5.2.
Sonority and the distribution of codas in English

Parallel to the previous discussion of place of articulation, this section determines whether sonorants are more frequent than obstruents in the rank order of English codas, based on the token word counts from the CDSC. The statistic used here was the Mann-Whitney U-Test for two independent samples, which is similar to the Kruskal-Wallis one-way analysis of variance, which is used for three or more independent samples. To compute the Mann-Whitney statistic, coda consonants were ranked according to their frequency (this

[4] This result held for each of the data sources:

Random House H (2, N=21)=.13, *p*=.94

Webster's H (2, N=21)=.01, *p*=.995

CDI H (2, N=21)=.67, *p*=.72

CDSC-type H (2, N=21)=.25, *p*=.88

Table 5 Rank of English codas based on token word counts from CDSC

Rank	CDSC token
21	t
20	r
19	n
18	d
17	z
16	k
15	s
14	l
13	m
12	v
11	ʃ
10	g
9	p
8	θ
7	ŋ
6	tʃ
5	f
4	dʒ
3	b
1.5	ʒ
1.5	ð

procedure was identical to that used for the Kruskal-Wallis test, see Table 5). Table 7 provides the ranks for sonorants and obstruents, along with the sum and mean rank for these groups.

This statistic determined whether within the rank order of English codas sonorants and obstruents form a homogeneous group, or whether the rank order these groups are distinct. The null hypothesis is that sonorants and obstruents are not reflected in the rank order of English codas. Results revealed that sonority was marginally reflected in the rank order of English codas, z=1.45, p=.07.[5] Although the results were not statistically significant, the pattern was in the direction one would predict based on cross-linguistic preferences in this position where sonorants are the preferred codas, given that in English the average rank of sonorants was higher than obstruents. Compare the mean rank of sonorants (14.6) to the mean rank of obstruents (9.88). This suggests that obstruent and sonorant codas exhibit different frequency characteristics. The question of whether children show evidence for these groupings is addressed in §5.4.2.2.

[5] The result of this test for each source is:

Random House z=1.86, p<.05

Webster' z=1.28, p=.1

CDI z=.87, p=.19

CDSC-type z=1.53, p=.06

The Random House Dictionary differs from the other sources in that all bimorphemic CVC words were excluded (see §3.2.1). To determine whether this pattern was unique, the type and token word counts from the CDSC were analysed with all bimorphemic words excluded from the corpus. Results were significant for CDSC-type (z=1.86, p<.05), and near significant for CDSC-token (z=1.61, p=.054).

Table 6 Kruskal-Wallis H Test applied to place of articulation in English coda ranks based on token word counts of the CDSC

Coronal		Dorsal		Labial	
t	21	k	16	m	13
r	20	g	10	v	12
n	19	ŋ	7	p	9
d	18			f	5
z	17			b	3
s	15				
l	14				
ʃ	11				
θ	8				
tʃ	6				
dʒ	4				
ʒ	1.5				
ð	1.5				
$n_1 = 13$		$n_2 = 3$		$n_3 = 5$	
Total = 156		Total = 33		Total = 42	
M coronal rank = 12		M dorsal rank = 11		M labial rank = 8.4	

3.5.3. Summary of place of articulation and sonority and the distribution of codas in English

An examination of place of articulation and sonority was needed to determine whether the rank order of English codas showed evidence for the most frequent place of articulation and the most frequent value of sonority (sonorants). Results showed that the rank order of occurrence of English codas does not reflect groupings by place of articulation. The data for sonority are less clear and suggest some evidence for sonority-based grouping in the English input. If child language data also show a grouping effect, then it will be difficult to distinguish the two hypotheses with respect to sonority. I will return to this issue in Chapter 5.

3.6. Summary

The current chapter provided a characterisation of the distribution of English coda consonants, upon which the SLGH is defined. The input-based hypothesis was formed using the rank order of English codas based on the token word counts of the CDSC because this database most closely approximates the phonological input for children acquiring English. It was also demonstrated that groupings based on place of articulation and sonority were not significantly reflected in the rank order of English codas; thus, the input-based hypothesis is argued to be best based on exemplars of codas. I now turn to child language data to determine whether children's productions more closely reflect universal preferences in coda position or the language-specific input.

Table 5 Rank of English codas based on token word counts from CDSC

Sonorants		Obstruents	
r	20	t	21
n	19	d	18
l	14	z	17
m	13	k	16
ŋ	7	s	15
		v	12
		ʃ	11
		g	10
		p	9
		θ	8
		tʃ	6
		f	5
		b	3
		dʒ	4
		ʒ	1.5
		ð	1.5
$n_1 = 5$		$n_2 = 16$	
Total = 73		Total = 158	
M sonorant rank = 14.6		M obstruent rank = 9.88	

CHAPTER 4
Child language codas

4.1.
Introduction

This chapter turns to child language acquisition to establish patterns of children's coda consonant production. The determination will be made through an examination of the coda consonants first produced by English-speaking children. The outline of the chapter is as follows. To begin, I describe the data on which the analyses are based. Following this, two methods are employed in the analysis of children's coda productions: Independent Analyses and Relational Analyses (to be discussed in §4.3).

4.2.
Data

Previous research looking at coda consonant acquisition has revealed that children's early coda productions are often obstruents (Fikkert, 1994; Goad, 1997a; Salidis & Johnson, 1997; Velten, 1943). More specifically, others have noted the special status in coda position of voiceless obstruents and nasals over voiced obstruents (Bernhardt & Stemberger, 1998; Kehoe & Stoel-Gammon, 2001). The research here explores whether the same patterns hold across data from a larger group of children. In fact, thus far, there has been no single comprehensive study on coda consonant acquisition in English. Part of the data analysed in this chapter is taken from a number of previously published studies whose findings and/or primary data have been coalesced to provide as clear a picture as possible on aspects of children's coda productions. There were a number of criteria for a study to be included in the evaluation. Data from previous studies are restricted to monolingual children acquiring English. For example, data from Leopold, 1939 and Velten, 1943 are excluded because the children in these studies were simultaneously acquiring languages other than English. In addition, data from children who were producing a variety of codas from the beginning of the study were excluded because these children were deemed to be too phonologically advanced for the purposes of this research. For this reason, data from Albright and Albright, 1956 were not examined, nor was data analysed from Smith, 1973 whose son Amahl produces all codas aside from /l/, continuants and voiced stops from the onset of the study—see Zamuner, 1996 for an analysis of Amahl's coda acquisition. Lastly, for the same reason, I have not included data from studies in which children were three years and older (e.g., Smit et al., 1990; Snow, 1963). A number of studies e.g., Menn, 1971 give both primary data and an analysis of children's coda productions. In many of these cases, the investigation focuses on a different aspect of coda production than assessed here. Thus, I evaluate only the primary data provided by the authors. All analyses when possible were restricted to CVC monosyllabic and monomorphemic words obtained from these sources. The motivation for this is that children differentially produce phonemes depending on their morphological status (Brown, 1973). Thus, an attempt was made to avoid plural and past tense morphemes to give a potentially less biased estimate of children's ability to produce particular phonemes, namely /d/ and /z/.

The other data analysed was based on two existing sources available from CHILDES: the Bernstein-Ratner database (Bernstein-Ratner, 1987), from which the data on Gina (1; 9–2; 1) and Lena (1; 7–2; 0) were taken; and the Higgenson database (Higgenson, 1985), which provided data from April (1; 10–2; 11) and June (1; 3–1; 9). These analyses on children's coda productions were restricted to CVC monosyllabic, monomorphemic words. A number of limitations were found in using these corpora. With respect to the Bernstein-Ratner database, children's phonetic transcriptions were not provided with a gloss and had to be determined from the transcription and the context. The only utterances included were ones in which the author and a second coder agreed on the targets that the children were attempting. Consequently, some words were not analysed because it was not clear what the child was saying. Although a gloss was provided with the Higgenson database, the phonetic transcription was continuous, which made it difficult at times to determine words' boundaries. Cases of ambiguity were not included in the analyses. The data and studies examined are provided in Table 1, along with the number of children in the study.

Table 1 Studies examined for data on coda consonant acquisition

References	Number of children or child's name
Bernstein-Ratner, 1994	9 children
Bernstein-Ratner Corpus (Bernstein-Ratner, 1987)	Gail
Bernstein-Ratner Corpus (Bernstein-Ratner, 1987)	Lena
Braine, 1974	Jonathan
French, 1989	Andrew
Higgenson, 1985	April
Higgenson, 1985	June
Holmes, 1927	Mollie
Ingram, 1974	Jennika
Massar (unpublished data)	Ben
Menn, 1974	Daniel
Ohala, 1992	Joey
Olmsted, 1971[6]	17 children
Salidis & Johnson, 1997	Kyle
Shibamoto & Olmsted, 1978	"F"
Stoel-Gammon, 1985[7]	19 children
Vihman, 1992; Vihman, Velleman & McCune, 1994; Vihman & McCune, 1994 (taken from Vihman 1996)	Alice
Vihman & Velleman, 1989 (taken from Vihman 1996)	Molly
Vihman, 1992; Vihman et al., 1994 (taken from Vihman, 1996)	Timmy

4.3. Independent and Relational Analyses

Following Stoel-Gammon (1985), two primary methods were used to analyse children's productions. The first is referred to as an "Independent Analysis." This analysis measures the codas children produce, without consideration of the target sounds children attempt. This "provides a complementary view of the child's phonological system by describing the sounds [...] produced regardless of their relation to the adult model" (Stoel-Gammon, 1985, p. 505). The second method of analysis is a "Relational Analysis," "which compares the child's pronunciation of a word with the adult standard form. The comparison may concentrate on the correct sound produced and describe the emergence and mastery of phonemes of the adult language" (Stoel-Gammon, 1985, p. 505). Thus, this method examines the accuracy of children's coda productions. Because the two analyses provide different information for how children acquire codas, I will first evaluate children's coda productions using an Independent Analysis (§4.3.1), then using a Relational Analysis (§4.3.2).

4.3.1. Independent Analysis of coda consonant acquisition

As mentioned above, an Independent Analysis measures each child's production of coda consonants without considering the target form. For example, Molly's production of "coat" as kʊk would be coded as a coda production of [k], with no evaluation of its incorrectness given the target is /t/. Using this evaluation method, the initial codas produced by children based on previous research and the CHILDES study are given in Table 2 and the results are summarised in Table 3. Given the variation inherent in looking at data across previously published research, the summary represents either the first codas produced by children (in these cases, either this was stated by the author or this was discerned by looking at the primary data) or the five most frequently produced codas (in these cases, this was the necessary measure because there was not a single coda listed or determined in my analyses as produced first).

Table 2 Independent Analysis of children's coda productions

Reference	Number of children or child's name	Age (in months)	First coda produced or 5 most produced codas
Bernstein-Ratner, 1994	9 children	13–25 M	n>t>k>d>m
Braine, 1974	Jonathan	19M	ʔ
French, 1989	Andrew	18M	tʃ
Higgenson, 1985	April	22–35M	t>n>s>k>p
Higgenson, 1985	June	15–21M	t>k>l>p>n
Holmes, 1927	Mollie	18M	k, t, d
Ingram, 1974	Jennika	15M	t
Massar	Ben	13M	t
Menn, 1974	Daniel	22M	r, F,[4] f, b~m

[6] A small number of parents in this study were not native speakers of English.

[7] These data were restricted to meaningful speech (if the child was producing at least 10 unique words). A sound was considered part of the child's inventory if it was produced in least two unique words.

Reference	Number of children or child's name	Age (in months)	First coda produced or 5 most produced codas
Ohala, 1992	Joey	18M	n, l, ʔ
Shibamoto & Olmsted, 1978	"F"	14M	t^h~t
Stoel-Gammon, 1985	19 children	18M	t
Vihman, 1992 and others	Alice	14M	ʔ
Vihman & Velleman, 1989	Molly	13M	ŋ, n, ʔ, t, tʃ
Vihman, 1992 and others	Timmy	11M	p

Table 3 Codas first produced by children, Independent Analysis

Coda	Number of children who initially produce coda
t	35
n	13
k	12
d	10
m	10
ʔ	4
P	3
l	2
tʃ	2
b	1
f	1
F	1
ŋ	1
r	1
s	1

Based on data from 41 children, children are most likely to produce the coda /t/. Coda productions based on these results is presented in Table 4, which gives the most frequent coda to the least frequent codas initially produced by children.

Table 4 Child language patterns for English codas: Independent Analysis

Child Language Patterns: Independent Analysis for English
t>n>k>d, m>ʔ>p>tʃ, l>b, f, F, ŋ, r, s

The Independent Analysis determined which codas children first produce without considering the target sounds. Now, the Relational Analyses of coda consonant acquisition for children acquiring English are presented.

4.3.2. Relational Analysis of coda consonant acquisition

The Relational Analysis, compares children's coda productions to the adult form. Using the same example, when Molly produces "coat" as kʊk this is coded as an error in the child's production of /t/, and would not be counted in a Relational Analysis. A summary of the data used for the Relational Analysis is given in Table 5, and results using this evaluation method are given in Table 6.

Table 5 Relational Analysis of children's coda productions

Reference	Number of children or child's name	Age (in months)	First coda produced or 5 most produced codas
French, 1989	Andrew	19M	s
Higgenson, 1985	April	22–35M	b
Higgenson, 1985	June	15–21M	f
Holmes, 1927	Mollie	13M	k, t, d
Ingram, 1974	Jennika	15M	t
Massar	Ben	13M	t
Menn, 1974	Daniel	22M	r
Ohala, 1992	Joey	18	n, l
Olmsted, 1971	17 children	15–23M	½ subjects correct—k[8]
Salidis & Johnson, 1997	Kyle	11M	d, k, m, p, s, t, z
Vihman, 1992 and others	Alice	14M	n
Vihman & Velleman, 1989	Molly	11M	ŋ, n, t

Table 6 Codas first produced by children, Relational Analysis

Coda	Number of children who initially produce coda
k	10.5
t	5
n	3
d	2
s	2
b	1
f	1
l	1
m	1
ŋ	1
P	1
r	1

[8] Correct is defined as at least 50% of the phoneme's productions in that position were produced correctly.

Coda	Number of children who initially produce coda
z	1

The summary in Table 7 presents the results obtained using a Relational Analysis, in order of the most frequent codas to the least frequent codas produced correctly. Results are based on data from 28 children.

Table 7 Child language patterns for English codas: Relational Analysis

Child Language Patterns: Relational Analysis for English
k>t>n>d, s>b, f, l, m, ŋ, p, r, z

4.4. Summary of child language coda preferences

The first method of children's acquisition of coda consonants was the Independent Analysis, which looked at children's coda productions independent of the adult target. The second method was the Relational Analysis, which looked at children's coda productions in relation to the adult target. There is both consistency and inconsistency found across the different analyses. Take for example the five most frequently produced codas as determined by the two analyses (Table 8).

Table 8 Five most frequently produced codas by children as established by Independent Analysis and Relational Analyses

Independent Analysis	Relational Analysis
t>n>k>d, m	k>t>n>d, s

Both results include the codas /t, n, k, d/; however, the Independent Analysis contains /m/ and the Relational Analysis contains /s/. Both results are therefore analysed separately with respect to the two hypotheses. Note that these results do not support the special status of fricatives in final position, as seen in Dinnsen (1996), Ferguson (1978) and Vihman and Ferguson (1987).

The next step was to evaluate children's coda productions with respect to the UGH established in Chapter 2 and the SLGH established in Chapter 3. These results are discussed at the end of Chapter 5, along with the results from an experiment looking at coda acquisition in English.

4.5. Limitations based on using data from previous studies

There are a number of limitations in using data from previous acquisition studies. I identify some caveats to consider when using existing data, which include problems in distinguishing between spontaneous versus imitated productions, prosodic environment, content versus function words, and type versus token frequencies. To begin, it has long been noted in the literature that children's spontaneous productions are less accurate than those produced as immediate imitations of adults (Olmsted, 1971); therefore, limitations arise if studies do not distinguish between these types of utterances. For example, as seen in Bernstein-Ratner (1994), patterns of results may differ if these different types of utterances are taken into account in analyses.

Another inadequacy in using previously published research is that studies often fail to control for the effects of prosodic position on children's productions, which is a factor that influences children's responses. For example, Zamuner and Gerken (1997) found that 2-year-old children's productions of codas were dependent on prosodic environment, such that both sonorant and obstruent codas were produced in stressed syllables, whereas children produced mainly sonorant codas in unstressed syllables. Learning to produce a coda in one position does not equate learning a coda across all positions. To circumvent these issues, acquisition studies must control for prosodic environment and separate these environments in analyses, yet few studies do so. For instance, Stoel-Gammon (1983) and Olmsted (1971) fail to control for this possibility.

When examining children's coda productions, another consideration is whether to collapse across word types, given that function words and content words have different phonological patterns and frequencies. For example, data obtained from the Higgenson corpus in CHILDES (as described in §4.2) were further analysed with respect to word type. When data are controlled for word type (e.g., content versus function words), the accuracy of children's coda productions changes. For example, the accuracy of April's production of /z/ decreases when function words are removed from the analysis. Given these type of results, it is appropriate and necessary to control for word class in analyses of children's productions.

Lastly, one of the major limitations in previous studies is that analyses and conclusions are drawn from a small data set. Acquisition is best characterised as probabilistic patterns seen across children. Thus, in order to draw conclusions that can be argued to hold across children, one needs sufficient data to support one's claims.

Take for example Salidis and Johnson (1997), whose goal was to evaluate a child's prosodic development with respect to the parameters proposed for prosodic development in Dutch (Fikkert, 1994). Because part of their interest was to assess Fikkert's claim that the first codas produced are obstruents, and that vowel length is first contrastive before sonorant codas, Salidis & Johnson analysed a single child's coda productions in terms of segmental class. According to their analyses, obstruents were more likely to be produced than sonorants in coda position; however, their finer-grained analysis found that the most accurately produce codas were plosives>nasals>fricatives>liquids, where obstruents and sonorants are interspersed. Furthermore, looking at individual segments, the most likely codas to be produced are /p, b, t, k, s, m/, and the most likely codas to be deleted are /g, θ, z, n, l, r/. Because of these limitations, it is difficult to determine the child's development at any given stage. For example, the child's accuracy in producing /d/ drops from 100% to 0% between 11 months and 13 months; however, this result is based on not more than two forms.

Further limitations that are encountered in using previously published research are that analyses do not necessarily weigh the same data equally. For example, a single child who produces "yes" as "yet" fifty times in a single session, would have a low accuracy on their production of /s/; however, this conclusion is primarily influenced by a single word type. In addition, many studies are based on data obtained from naturalistic settings and not all coda types are represented in the samples; therefore, the conclusions drawn are based on a subset of the possible data.

In summary, a number of important factors have been identified as relevant for analyses of language acquisition: spontaneous versus imitated productions, prosodic environment, content versus function words, restricted data sets, and type versus token frequencies. These limitations illustrate the importance of the experiment described in Chapter 5, which is taken as a starting point to a comprehensive examination of coda consonant acquisition in English.

CHAPTER 5
Experiment 1
Two-year-olds' production of coda consonants

5.1. Introduction

This chapter presents results from an experiment examining the production of coda consonants in children acquiring English. Specifically, the experiment was designed to test children's production of a variety of codas, which allows for a more complete picture of coda consonant acquisition. These data are missing from previous naturalistic studies because there has been no comprehensive study on coda consonant acquisition in English. The goal was to present children with real words containing a variety of codas. There are a number of strengths to the present experiment. First, it controls for children's spontaneous and imitated productions. In addition, the set of words tested were restricted to CVC content words, which avoids any confounds that may arise from collapsing prosodic environment and word class. Lastly, in the present study 17 children were tested, allowing for the results to be collapsed across children and allowing for a more general determination of children's acquisition of coda consonants in English.

Following a presentation of the experiment design and results, I then turn to an evaluation of the UGH and the SLGH with respect to the results of the experiment, and to the results of the Independent Analysis and the Relational Analysis from Chapter 4.

5.2. Method

5.2.1. Subjects

Subjects included in the analyses were 17 children between the ages of 20 and 26 months (M=25.5 months), 10 males and 7 females, with a mean MLU of 1.92. All children were monolingual speakers of English with no history of speech or hearing impairment as determined by parent questionnaire (3 subjects were included who were exposed to Spanish for no more than 10 hours a week). All parents completed the MacArthur Communicative Development Inventory (CDI) (Fenson et al., 1993), which has been shown to be an accurate predictor of children's later language development. Subjects had to achieve above the 10^{th} percentile on Part A: Vocabulary Checklist (range 10–95th percentile, M=52.5). A subset of the participants also completed the experiment described in Chapter 6.

Not included in the analyses were 12 subjects that were tested but failed to complete the first portion of the experiment (see description of stimuli §5.2.2). These subjects were not yet producing many words or they were not interested in participating in the task. It is assumed that these data are missing at random and that these subjects' performance on the task was unrelated to the purposes of the experiment. Lastly, another 5 subjects were excluded: 2 subjects were bilingual, 2 subjects had a history of hearing and language impairment, and 1 subject failed to reach the 10th percentile on the Vocabulary Checklist of the CDI.

The rationale for including subjects just under and over 2 years of age is that this is when children are both producing and deleting codas. An attempt was made to include subjects as young as possible (20 months) to determine whether their performance would differ from the older children's performance (27–26 months).[9] Most subjects, however, were between 26 and 27 months of age; thus, additional younger subjects are needed to further explore this possibility. All subjects were recruited through the Language Acquisition lab at the University of Arizona.

5.2.2. Stimuli

The stimuli consisted of 70 English monomorphemic CVC[10] words. Each of the 70 words had a corresponding picture stimulus that was scanned from Richard Scarry's "Best word book ever" or "Busy, busy town." All except 14 words appeared in the CDI, which contains vocabulary items of common words known to children. These fourteen words were: "witch," "laugh," "king," "shell," "buzz," "tub," "bell," "wing," "drum," "tail," "five," boat," "road," and "watch." Of these words, only "buzz" did not occur in the CDSC, described in §3.2.4. Parents completed a questionnaire to determine children's familiarity with these words. Parents' responses are summarised in Table 1. All items were identified as familiar to the children by at least one quarter of all parents.

Table 1 Familiarity of words not present on CDI

Identified as familiar	Word
by ¼ of parents	shell, king
by ½ of parents	laugh, buzz, five, scratch, drum, bell, wing, road
by ¾ of parents	tail

The words were presented in 4 blocks. Each block contained unique words ending in the set of possible codas in English (/tʃ, g, b, r, p, f, ŋ, d, k, s, v, n, t, l, z, ʃ, m, θ/), except /ð/, /ʒ/ and /dʒ/ which are highly restricted in this position, and therefore were not tested. Vowel length was controlled for in each block. Thus, the first two blocks contained only words with lax vowels, while words in the third and fourth block contained tense vowels. The decision to control for vowel quality (lax or tense) was based upon previous research (Fikkert, 1994) showing that children differentially produce codas depending on the quality of the preceding vowels. The list of words is given in Table 2. The experiment had a within-subjects design and all subjects were tested on the same stimuli.

[9] Younger children aged 20–24 months were slightly less accurate in their coda productions (M=9 correct codas produced) versus older children aged 26–28 months (M=10.7 correct codas produced). In addition, younger children produced fewer words spontaneously (M=3) than older children (M=4.8).

[10] Seven words had complex onsets: "crib," "brush," "scratch," "drum," "slide," "stove" and "frog."

5.2.3. Procedure

Subjects were tested in the Language Acquisition Lab at the University of Arizona with the exception of 1 subject, who was tested at home. Subjects tested at the University of Arizona were videotaped and the subject tested at home was audio-recorded using a DAT recorder.

Subjects were shown various pictures presented on a computer. Responses were elicited by engaging children in the pictures and by asking subjects to identify specific animals or objects. Often children produced the target word spontaneously; if not, the experimenter would either produce the word (which the child would often repeat without being prompted) or the experimenter explicitly asked the child to repeat the specific word.

5.3. Results

5.3.1. Data coding

Data were transcribed according to the standard practice used in phonological acquisition research. All sessions were transcribed on-line and checked by a research assistant naïve to the experiment's design. Reliability between the two transcribers was determined by a third person, who was also naïve to the purposes of the experiment, and calculated as 99.3% accurate (reliability was based on 23.5% of the data). No acoustic measurements were performed on children's responses; however, future analyses will examine whether children made covert distinctions in their productions that were not audible to the transcribers. Children's first response was coded for whether the word was produced spontaneously (S) or as an imitation (I). Responses were coded as imitated if the word was produced by an adult prior to the child's production of that word during the experiment session. Responses were also coded for whether children accurately produced the coda consonant.

5.3.2. Results of spontaneous and imitated productions

There were too few responses for words in blocks two to four of the stimuli to analyse; therefore, the analyses focus on data only from the first block. The blocks were presented in a set order (blocks 1 to 4), and many children lost interest after completing the first block. The results of children's spontaneous and imitated production of codas in words from the first block are given in Table 3.

As seen in Table 3, overall there were fewer spontaneous than imitated productions. In addition, certain words were never produced spontaneously, such as "witch" and "kiss." Future experiments would benefit from having the full range of codas produced both spontaneously and as imitations to allow for a complete analysis. From the data in Table 3, one can see that children were more accurate in producing codas when they were produced spontaneously, rather than when they were produced as imitations (72% versus 64%, respectively (χ^2 (1, N=184)=23.67, p<.001)). This result differs from previous results in the literature that has found children are more phonologically accurate in imitated productions (Olmsted, 1971). This contrast can be attributed to differences in experiment design. For example, data from Olmsted's study were collected in naturalistic settings, so all items were presumably familiar. In this study, however, the words

Table 2 Stimuli (Lax=Lax Vowels, Tense=Tense Vowels)

Coda	Block 1–Lax		Block 2–Lax		Block 3–Tense		Block 4–Tense	
tʃ	witch	wɪtʃ	scratch	skrætʃ	couch	kaʊtʃ	watch	watʃ
g	leg	lɛg	pig	pɪg	dog	dag	frog	frag
b	crib	krɪb	tub	tʌb	–	–	–	–
r	bear	bɛr	hair	hɛr	car	kar	door	dor
p	cup	kʌp	lip	lɪp	soap	sop	sheep	ʃip
							soup	sup
f	laugh	læf	foot	fʊt	knife	naɪf	roof	ruf
ŋ	king	kɪŋ	wing	wɪŋ	–	–	–	–
			sing	sɪŋ				
d	bed	bɛd	head	hɛd	slide	slaɪd	food	fud
							road	rod
k	book	bʊk	duck	dʌk	cake	kek	sock	sak
s	kiss	kɪs	bus	bʌs	house	haʊs	mouse	maʊs
							juice	dʒus
v	love	lʌv	give	gɪv	five	faɪv	stove	stov
n	sun	sʌn	pen	pɛn	moon	mun	rain	ren
t	cat	kæt	hat	hæt	boat	bot	coat	kot
l	shell	ʃɛl	bell	bɛl	tail	tel	ball	bal
z	buzz	bʌz	–	–	–	–	cheese	tʃiz
ʃ	fish	fɪʃ	brush	brʌʃ	wash	waʃ	–	–
			dish	dɪʃ				
m	swim	swɪm	lamb	læm	comb	kom	–	–
			drum	drʌm				
θ	bath	bæθ	–	–	tooth	tuθ	mouth	maʊθ

that children spontaneously produced were words familiar to children, whereas words that were imitated were probably not as familiar. Given that children's spontaneous productions of words are a more accurate measure of what children know, the remaining analyses are restricted to children's spontaneous productions.[11]

Table 3 Children's coda responses on English real-word stimuli

Coda	Word	Spontaneous Coda Productions	Imitated Coda Productions
tʃ	witch	NR*	11/15 (73.3%)
g	leg	0/2 (0%)	7/12 (58.3%)
b	crib	2/2 (100%)	8/13 (61.5%)
r	bear	10/11 (90.9%)	4/5 (80%)
p	cup	5/6 (83.3%)	9/9 (100%)
f	laugh	NR	10/14 (71.4%)
ŋ	king	0/1 (0%)	6/14 (42.9%)
d	bed	7/8 (87.5%)	4/8 (50%)
k	book	8/9 (88.9%)	7/8 (87.5%)
s	kiss	NR	12/14 (85.7%)
v	love	NR	9/12 (75%)
n	sun	8/9 (88.9%)	5/5 (100%)
t	cat	8/9 (88.9%)	5/7 (71.4%)
l	shell	NR	8/14 (57.1%)
z	buzz	NR	5/10 (50%)
ʃ	fish	6/13 (46.2%)	2/6 (33.3%)
m	swim	3/3 (100%)	10/11 (90.9%)
θ	bath	1/8 (12.5%)	1/6 (16.7%)
Total		59/82 (72%)	125/196 (64%)

*NR=no responses.

As seen in Table 3, overall there were fewer spontaneous than imitated productions. In addition, certain words were never produced spontaneously, such as "witch" and "kiss." Future experiments would benefit from having the full range of codas produced both spontaneously and as imitations to allow for a complete analysis. From the data in Table 3, one can see that children were more accurate in producing codas when they were produced spontaneously, rather than when they were produced as imitations (72% versus 64%, respectively (χ^2 (1, N=184)=23.67, p<.001)). This result differs from previous results in the literature that has found children are more phonologically accurate in imitated productions (Olmsted, 1971). This contrast can be attributed to differences in experiment design. For example, data from Olmsted's study were collected in naturalistic settings, so all items were presumably familiar. In this study, however, the words that children spontaneously produced were words familiar to children, whereas words that were imitated were probably not as familiar. Given that children's spontaneous productions of words are a more accurate measure of what children know, the remaining analyses are restricted to children's spontaneous productions.[12]

[11] Children's spontaneous coda productions were coded to determine if when children misarticulated a coda, whether the substituted coda was more or less frequent. Results showed that out of 26 coda errors, 24 of these errors were when a less frequent coda was substituted with a more frequent coda. For example, /g/, /f/ and /ʃ/ were substituted as [k], [p] and [s], respectively.

5.3.3.
Results based on weighted responses of spontaneous productions

Because there were unequal numbers in children's spontaneous productions, a weighted statistic of children's correct coda responses was calculated. It was necessary to calculate responses in this fashion because not all experimental stimuli were equally produced spontaneously. As seen in (1), the weighted statistic was computed by weighting the number of words produced with the coda (Nw) divided by the number of times the word with the coda could have been produced (17). (There were 17 opportunities for each word to be spontaneously produced because there were 17 subjects in the study.) This proportion was then multiplied by the number of correctly produced codas (NCPD).

(1) Weighted Responses $= \frac{N_W}{17} * N_{CPC}$

To illustrate, the word "bear" was produced spontaneously 11 out of 17 times (.647 of the time). To calculate a weighted response, .647 was multiplied by the number of times the coda in the word "bear" was produced accurately (.647*10 =6.47) (also see (2)). Children's coda responses based on the weighted statistic are given in Table 4 below. Items that children did not produce spontaneously were not included in the analysis ("witch," "laugh," "kiss," "love," "shell," and "buzz").

(2) Weighted Response (bear) $= \frac{11}{17} * 10$
$= 6.47$

Table 4 Weighted responses (WR) for spontaneous coda productions

Coda	Word	WR for Spontaneous Coda Productions
g	leg	0
b	crib	0.24
r	bear	6.47
P	cup	1.77
ŋ	king	0
d	bed	3.29
k	book	4.24
n	sun	4.24
t	cat	4.24
ʃ	fish	4.59
m	swim	0.53
θ	bath	0.41

All subsequent analyses and discussions concerning the results of this experiment are based on the weighted statistic. The accuracy of children's coda productions in this experiment can be summarised from the most accurate coda produced to the least accurate coda produced in Table 5 below.

[12] Children's spontaneous coda productions were coded to determine if when children misarticulated a coda, whether the substituted coda was more or less frequent. Results showed that out of 26 coda errors, 24 of these errors were when a less frequent coda was substituted with a more frequent coda. For example, /g/, /f/ and /ʃ/ were substituted as [k], [p] and [s], respectively.

Table 5 Child language patterns for English codas: Experiment 1

Child Language Patterns: Experiment Analysis for English
r>ʃ>t, n, k>d>p>m>θ>b>ŋ, g

Note that /r/ and /ʃ/ were the most accurately produced codas in this experiment. Presumably this reflects children's familiarity with the words "bear" and "fish," and increases their accuracy in producing these codas. Further experiments would benefit from having and need to have a range of word types to test children's coda acquisition.

5.4. Evaluation of the UGH and the SLGH

It is now possible to evaluate the UGH and the SLGH with respect to data from children acquiring English. Before turning to these analyses, a summary will be provided for both hypotheses and for how the predictions were formulated. In addition, a synopsis will be provided for the child language data obtained in the experiment described in this chapter, and from the Independent and Relational Analyses described in Chapter 4.

5.4.1. Summary of UGH, SLGH and child language data

5.4.1.1. Summary of the UGH

Recall the UGH for coda consonant acquisition was established according to previously published research and data from lexicons of CVC words from 35 languages. Results showed that coronal place of articulation is preferred in coda position and that languages prefer codas that are sonorant (see Table 6). The interpretation of UG considered in this dissertation is that cross-linguistic patterns reflect UG. Thus, if language acquisition is mediated by innate properties of UG, the prediction is that children will prefer coronal and sonorant codas. The UGH for coda consonant acquisition is restated in (3).

(3) Universal Grammar Hypothesis (UGH2): Children will produce coronal codas (t, d, s, z, n, l, r) more than labial or dorsal codas, and sonorant codas (m, n, ŋ, l, r) more than obstruent codas in word-final position.

These predictions are taken as an initial point for a UG hypothesis of coda consonant acquisition. Realistically, one would not predict children to initially produce segments such as /tʃ/ and /dʒ/ in coda position, given that they are marked for other reasons. The approach taken here is to determine whether children's initial coda productions bear coronal place of articulation over labial or dorsal place of articulation, a pattern predicted by the UGH. Thus, the expectation is that children will initially produce the codas /t/, /d/, /n/ over codas such as /p/ and /k/.

5.4.1.2.
Summary of the SLGH

Chapter 3 established the predictions of the SLGH for coda consonant acquisition based on the token word counts of the CDSC. This corpus was argued to be the closest approximation to the language-specific input for children acquiring English. In addition, previous research has illustrated that acquisition models based on token word counts are more accurate than those based on type word counts. Also, although it is possible to base predictions for acquisition on the frequency of codas in English, the choice was made to predict acquisition on the rank order of codas. The result of the English ranking is given in Table 6 and the input-based hypothesis for children's acquisition of coda consonants is restated in (4).

(4) Specific Language Grammar Hypothesis (SLGH): Children will produce the most frequent codas in word-final position: t>r>n>d> z>k>s>l>m>v>ʃ>g>p>θ>ŋ>tʃ>f>dʒ>b>ʒ, ð.

The UGH and the SLGH are similar in that they both predict children to show preferences for codas with coronal place of articulation. They differ, however, in that the SLGH predicts children will also initially produce the frequent coda /k/, despite it having dorsal place of articulation. With respect to sonority, the UGH predicts that children will display a preference for sonorant codas, whereas the input-based model of acquisition predicts preferences for both sonorants and obstruents. Arguments were made that the predictions of the SLGH should be based on the rank order of individual English coda consonants, rather than on groupings based on place of articulation and sonority. See §3.5 for this discussion.

Table 6 Coda consonants: X-Linguistic, Language Specific Patterns

X-Linguistic patterns	Language Specific Patterns: English
t, d, s, z, n, l, r (ʃ, ʒ, tʃ, dʒ, θ, ð)	t>r>n>d>z>k>s>l> m>v>ʃ>g>p>θ>ŋ>tʃ>f> dʒ>b>ʒ, ð.
m, n, ŋ, l, r	

5.4.1.3.
Summary of child language data

Children's production of coda consonants in English were characterised in three ways. First, data were coalesced from a number of previous published studies and from data from CHILDES described in Chapter 4. These data established the codas children initially produce (Independent Analysis) and the codas children initially produce accurately (Relational Analysis). Lastly, data were obtained from an experiment testing children's production of a variety of codas. The results of these studies are given in Table 7.

Table 7 Child Language patterns English: Independent Analysis, Relational Analysis and Experiment 1

Analysis	Child Language Patterns: English
Independent Analysis	t>n>k>d, m>ʔ>p>tʃ, l> b, f, F, ŋ, r, s
Relational Analysis	k>t>n>d, s>b, f, l, m, ŋ, p, r, z
Experiment 1	r>ʃ>t, n, k>d>p>m>θ>b>ŋ, g

Both the UGH and the SLGH are now evaluated with respect to patterns established for children's acquisition of coda consonants in English.

5.4.2. UGH and child language data

The analyses in this section determine whether children's production of coda consonants mirrored the UGH, which predicts that children will prefer codas that have coronal place of articulation, as well as codas that are sonorant over obstruent (see Tables 6 and 7 above). It was first necessary to establish a means by which children's coda productions could be evaluated according to the UGH. The approach taken here was to measure whether coronal codas and sonorant codas were the most frequently produced codas by children acquiring English (see § 3.5 for a further discussion of the statistic used).

5.4.2.1. UGH and place of articulation

To determine whether coronal codas were preferred in children's production of codas, the Kruskal-Wallis one-way analysis of variance was computed. Children's coda consonant productions (determined by various means), were ranked according to their frequency. The test measured whether the groups of coronal, dorsal and labial codas were from the same population. The null hypothesis is that place of articulation is not reflected in the rank order of codas children first produce. The total and average ranks for coronal, dorsal and labial codas for each child language analysis are given in Table 8, and results are presented in Table 9.

Table 8 Total and average ranks for coronal, dorsal and labial codas for Independent Analysis, Relational Analysis and Experiment 1

	Independent[13]			Relational			Experiment 1		
	Cor	Dor	Lab	Cor	Dor	Lab	Cor	Dor	Lab
n	9	2	4	7	2	4	6	3	3
Total	76	16.5	27.5	55.5	17.5	18	52	12	14
M rank	8.44	8.25	6.88	7.93	8.75	4.5	8.67	4	4.67

Table 9 Kruskal-Wallis results for place of articulation in children's coda productions

Analysis	Result
Independent Analysis	H (2, N=15)=.35, p=.84
Relational Analysis	H (2, N=13)=2.45, p=.29
Experiment 1	H (2, N=12)=4.38, p=.11

Results found no evidence that the average rank of coronal codas was significantly higher or distinct from that of dorsal and labial codas.

5.4.2.2. UGH and sonority

The second set of analyses determined whether children's coda productions were more frequently sonorants or obstruents. These analyses are parallel to the previous analyses of place of articulation. The statistic used for the evaluation was the Mann-Whitney U-Test for two independent samples. To compute the Mann-Whitney statistic, children's productions of coda consonants were ranked according to their frequency.

Table 10 provides the ranks for sonorants and obstruents, along with the mean rank for these groups. The results are given in Table 11.

Table 10 Total and average ranks for sonorant and obstruent codas for Independent Analysis, Relational Analysis and Experiment 1

	Independent		Relational		Experiment 1	
	Son	Obs	Son	Obs	Son	Obs
n	5	10	5	8	4	8
Total	40	80	29	62	27.5	50.5
M rank	8	8	5.8	7.75	6.88	6.31

Table 11 Mann-Whitney U-Test results for sonority in children's coda productions

Analysis	Result
Independent Analysis	z=0, *p*=.5
Relational Analysis	z=.81, *p*=.21
Experiment 1	z=.17, *p*=.43

Results showed no evidence for sonorants and obstruents forming distinct groups within the rank order of English codas produced by children. Moreover, not only were the patterns not significant, but they also varied across the different analyses of children's productions. With the Independent Analysis, the average rank for sonorants and obstruents was equal, whereas the Relational Analysis had a higher average rank for obstruents. Finally, the experimental results had a slightly higher average rank for sonorant codas. Again, there was no evidence for sonority-based groupings.

5.4.2.3. Summary of UGH evaluation

To evaluate child language with respect to the UGH, the approach was to see whether children's initial coda productions were more likely to be coronal and whether they were more likely to be sonorants. The present analysis did not yield evidence that these codas were preferred in children's coda productions. For example, compare the five most frequently produced codas as established by the Independent Analysis, the Relational Analysis and Experiment 1 (Table 12).

Table 12 Five most frequently produced codas by children as established by Independent Analysis, Relational Analysis and Experiment 1

Analysis	5 most frequently produced codas
Independent Analysis	t>n>k>d, m
Relational Analysis	k>t>n>d, s
Experiment 1	r>ʃ>t>n>k

[13] The segments /ʔ/ and /F/ (unspecified fricative) were analysed as coronals, under the assumption that these segments are underspecified for place of articulation and, therefore, would pattern with coronals. The analyses were also done with these segments excluded, with the same results.

The UGH predicted that children's initial coda productions would be coronal and sonorants. The results in Table 12 illustrate that although coronals and sonorants are among the set of most frequently produced codas, there are also other codas produced that are not predicted by the UGH, namely /k/ (and in one case /m/). The production of /k/ in coda position cannot be accounted for under the current interpretation of UG; however, a frequency analysis can offer an explanation for this pattern—that is, /k/ is the sixth most frequent coda in the input.

It is worth stating again that the characterisation and evaluation of UG offered in this dissertation is not exhaustive, but rather it is taken as a starting point for further research. UG is often given as an account for the patterns seen in children's early language productions and as a means to explain how children's phonological systems develop. The work here illustrates the complexity faced in predicting specific acquisition patterns from this model and in applying this model to existing data from child language. It is possible that other interpretations of the UGH may be able to account for the presence of /k/. Child language data are now evaluated with respect to the SLGH, which predicts that children will initially produce the most frequent codas in English.

5.4.3. SLGH and child language data

Children's coda productions were then analysed with respect to the SLGH, which predicts that children will initially produce the most frequent codas in English. A difference in the predictions made by the UGH and the SLGH is that only the SLGH predicts a significant correlation between the frequency of codas in the input and children's production of coda consonants. Children's coda productions from the Independent Analysis, the Relational Analysis and the experiment presented in this chapter were correlated with the frequency of English codas as determined by the frequency of codas in CDSC. Correlations allow one to look at the statistical relationship between two variables. Children's coda productions established through a variety of means were all significantly correlated to the frequency of codas in English (see Table 13).

Table 13 Correlations between children's coda productions and the frequency of codas in English

Analysis	Correlation (one tailed)
Independent Analysis	r_s=+.51, n=13, p<.05
Relational Analysis	r_s=+.59, n=11, p<.05
Experiment 1	r=+.8, n=10, p<.01

One possibility for why the SLGH appears to better account for the acquisition data than the UGH is that the analyses of the UGH might be more stringent. A analogous analysis of the SLGH was done where frequency was treated as a category. Recall that the SLGH predicted that children will prefer coronal codas (13 in English) over codas with dorsal and labial (8 in English) place of articulation. The SLGH was then tested by determining whether children preferred the most frequent codas in their productions over the less frequent codas. Results of the Mann-Whitney U-Test showed significant effects in two of the three operalisations of child language data and a marginal effect in the third: Independent Analysis (z=1.85, p<.05),[14] Relational Analysis (z=1.18, p=.12), Experiment 1 (z=1.94, p<.05). The UGH was also tested by determining whether children show evidence for the groupings of sonorants (5 in English) and obstruents (16 in English). The equivalent test with the SLGH is whether children prefer the 5 most frequent codas over the 16 less frequent codas in English. Results of the Mann-Whitney U-Test also showed marginal significance across the three operalisations of child language data: Independent Analysis (z=1.39, *p*=.08), Relational Analysis (z=.88, *p*=.19), Experiment 1 (z=1.78, *p*<.05). Therefore, even though the UGH might be more stringent, it is unlikely that this can account for its failure to accurately predict children's acquisition of codas.

5.4.4. Summary of UGH and SLGH evaluations

The primary question asked in the dissertation is whether language acquisition is best characterised as the unfolding of highly specified innate abilities (UGH) or whether language acquisition is best characterised as a domain-neutral learning process in which frequently occurring properties of the input serve to organise children's linguistic representations (SLGH). The hypotheses are contrasted through an examination of young children's acquisition of coda consonants in English.

The first view (UGH), that language is acquired based on innate properties of language, is traditionally based on cross-linguistic markedness observations from adult languages. To this end, the cross-linguistic preferences for place of articulation and sonority in coda position were determined through an examination of previously published research, and my own examination of data from 35 languages (Chapter 2). The cross-linguistic patterns established here were argued as the basis for the UGH for coda consonant acquisition.

The second view (SLGH) is that language is acquired based on the patterns in the input language. This is established, in part, on the large body of research illustrating that children's language reflects their specific language input at an early age. This hypothesis is an alternative to the UGH, given that the properties of language that are argued to be unmarked (and innate) are also often the most frequent properties in any one language. Thus, although one could argue that children initially produce unmarked structures due to innate UG, one could equally argue that children initially produce unmarked structures because these patterns are the most frequent.

To evaluate these theories, child language data were obtained from a number of sources: from previously published research, from an examination of available data from the CHILDES database (Independent Analysis and Relational Analysis), and from an experiment looking at children's production of a variety of English coda consonants. Children's coda productions from these different sources were then evaluated with respect to the UGH and the SLGH. The results of these analyses are summarised in Table 14. "✘"represent no relationship between the predicted and attested patterns in acquisition. "✓" indicates that the predicted patterns were mirrored in children's production of coda consonants.

Table 14 Children's coda productions in English from different sources and how they relate to the UGH and SLGH

Analysis	UGH	SLGH
Independent Analysis	✘	✓
Relational Analysis	✘	✓
Experiment 1	✘	✓

[14] The segments /ʔ/ and /F/ (unspecified fricative) were excluded from the Independent Analyses as they do not occur in the English input as phonemes.

The UGH was evaluated by determining whether there was evidence in children's coda productions for coronal, dorsal and labial place of articulation, and for sonorants and obstruents. Place of articulation was analysed using the Kruskal-Wallis one-way analysis of variance. Results found no evidence that the average rank of coronal codas was significantly higher or distinct from that of dorsal and labial codas. Similarly, results of the Mann-Whitney U-Test for two independent samples showed no evidence for sonorants and obstruents forming distinct groups within the rank order of English codas produced by children.

The SLGH was evaluated by correlating children's coda productions with the frequency of English codas (as determined by the frequency of codas in the CDSC). Result indicated that children's coda productions pattern with the frequency of codas in the ambient language, given the significant relationships between child data and the distribution of codas in the ambient language. These results held across the first codas children produce (Independent Analysis), the first codas children produce correctly (Relational Analysis), and based on data obtained from an experiment testing children's production of a variety of codas in English.

Previous research by Bernstein-Ratner (1994) looked at the frequency of codas in the input with respect to children's coda productions, yet no relationship was found. (No statistics were provided for the analyses.) In her study, children's final consonant productions at the preverbal, one-word stage and multi-word stage were compared to the maternal input. One limitation of this study is that the frequency of codas was collapsed across prosodic positions and word types. This can drastically change findings, given that children's productions of codas can differ based on these factors. In addition, Bernstein-Ratner's study used a phonetic analysis to convert the adult orthographic representations to phonetics, and from this the frequency of sounds in the adult input was calculated. These reasons might account for why no relationship was found between the frequency of codas in the input and children's production of codas.

Other work by Stoel-Gammon (1995, cited in Stoel-Gammon, 1998), has found various effects depending on type of analysis. Her study tested whether the more frequent phonological features are acquired early and less prone to errors in production. Results were mixed, and most strikingly not supported for final fricatives which are frequent in the input, yet as a class, are acquired late. The difficulty in a feature analysis, however, is that although the class of fricatives might be acquired late, certain fricatives are acquired relatively early. Stoel-Gammon also correlated the frequency of codas based on the type counts of words from the MacArthur Communicative Development Inventory (CDI) (Fenson et al, 1993) to the accuracy of 3-year-olds coda productions from Templin (1957). At the phonemic level, a significant correlation was found between the frequency of occurrence in target words and the accuracy of children's productions.

5.5.
Influences of methodology on results

There are a number of possible influences of the particular methodological factors used on the results of the present experiment, and the results presented here are taken as an initial indication of patterns that might hold across acquisition in English. Addressing these factors in future studies will help determine the generality of the effect. To begin, the study is restricted in that it considers coda consonants from single word types, e.g., children's production of /r/ is based on the single word "bear." Thus, it would be beneficial to have acquisition data from codas in different word types, such as children's productions of /r/ in the words "bear," "deer," and "door."

Another possible influence of methodology on the results of the present experiment is that the stimuli were not controlled for the characteristics of the surrounding consonants, such as the onset's place of articulation. This is potentially problematic because of phonological processes seen in child language of fronting and backing. Children routinely substitute consonants' places of articulation based on the other consonants in the word (Vihman, 1996). Thus, children's spontaneous responses in this experiment were evaluated for whether the onset's place of articulation could be responsible for children's misarticulated codas. In fact there were only three errors of this type for all of the responses (Table 15).

Table 15 Place of articulation errors in children's spontaneous coda productions linked to place of articulation in onset

Gloss	Adult Target	Child's production	Coda Error
cup	kʌp	tʌk	p→k
bath	bæθ	bæf	θ→f
bath	bæθ	bæf	θ→f

Given that there are so few examples of fronting and backing in children's responses, it seems reasonable to conclude that this factor did not significantly effect the pattern of results obtained in this experiment. Children's coda productions in this experiment were not underestimated due to difficulties children had in producing words that had with different places of articulation for the onsets and codas. It is also possible to examine children's coda production in relation to their onset production (as done in Coberly, 1985). This would allow one to determine whether child are able to generalise phonemes across environments.

Although the analyses in this chapter contrasted children's productions of different codas to the frequency of codas in English with positive results, it is also possible that children's productions of the same codas differ according to the frequency of surrounding sounds in the word. This possibility was tested in an experiment presented in the following chapter. In this experiment, a deeper interpretation of the input is explored, which is one of phonotactic probability. The goal of the dissertation is not just to look at which codas more frequently correlate with children's productions, but to also examine specific factors of the input that might contribute to the identical coda produced differently in the words of a language.

CHAPTER 6
Experiment 2
Two-year-olds' production of coda consonants andphonotactic probability

6.1.
Introduction

The research in this dissertation has examined children's acquisition of different coda consonants with respect to cross-language preferences (UGH) and the distribution of codas in English (SLGH). The research in this chapter seeks to further explore the input-based hypothesis of coda acquisition by providing a detailed examination of the role of the input in children's acquisition of coda consonants. Specifically, the goal is to investigate children's coda productions in different probabilistic contexts (discussed below). The hypothesis is that phonotactic probability is a factor that can account for children's varied production of the identical coda consonant.

6.1.1.
Phonotactic probability

Phonotactic probability concerns the likelihood of sounds' occurrences. For example, in English there are absolute restrictions where /ŋ/ can only occur in syllable-final position, such as in "wing" /wɪŋ/, and /h/ cannot occur in syllable-final position. There are also less absolute restrictions, e.g., /ʒ/ occurs in a total of three CVC words in the Webster's Lexicon: "beige" "rouge" and "loge," thus /ʒ/ has a very low probability of occurrence in this position. The probability of an event *(e)* is "computed as the relative frequency with which *e* occurs [n_e] in a sequence of *n* identical experiments [n]" (Bod, 2001).

Given that there are 3 instances of /ʒ/ out of a total number of 1,267 CVC words in the Webster's Lexicon, the probability of /ʒ/ occurring at the end of an English CVC word based on this dictionary is .0024 (1).

(1)

$$P\,(ʒ) = \frac{3}{1{,}267}$$

$$P\,(ʒ) = .0024$$

Other codas, however, occur at the end of many CVC words, e.g., /t/ as in "fight" "hat" "hit" "net" "lout" "shoot"; thus, the probability of /t/ is higher, and calculated to be. 16(2).

(2)

$$P\,(t) = \frac{204}{1{,}267}$$

$$P\,(t) =.16$$

Phonotactic probability can be calculated according to segmental positional probabilities and biphoneme probabilities. Segment positional probabilities refers to the likelihood of phonemes occurring in the onset, vowel and coda position (C-V-C), as in the example above. Biphoneme probabilities refers to the likelihood that phonemes are preceded or followed by other phonemes, that is, the likelihood of an onset-vowel sequence and a vowel-coda sequence (CV and VC, respectively).

6.1.2. Previous research on phonotactic probabilities in English

There is a large body of research looking at phonotactic probability in English CVC words. I restrict my review to research looking at just CVC words, although considerable work has been done on different word types (e.g., Messer, 1967; Hammond, 2001; Pierrehumbert, 1994; Pierrehumbert & Coleman, 1997). Although these studies looking at CVC probabilities used different tasks and spanned age groups, they have found that infants, young children, and adults are sensitive to phonotactic probability. Specifically, all groups show a preference for non-words composed of high phonotactic probabilities. For example, Jusczyk, Luce and Charles-Luce (1994) presented infants with novel CVC words with either low or high phonotactic probabilities and found that by 9 months, infants showed a preference for the high phonotactic probability non-words. This result demonstrates that even at an early age and before the onset of meaningful speech, infants are sensitive to the frequency of sound patterns. It also demonstrates that they are able to encode this information at some level. Other studies, such as that by Treiman, Kessler, Knewasser, Tincoff and Bowman (2000) revealed that subjects maintained the integrity of frequent VC rhymes more than infrequent VC rhymes in a non-word blending task. Similar sensitivities are illustrated in a study by Storkel (1999, 2001), which found that children learn high phonotactic probability non-words with fewer exposures than non-words with low probabilistic phonotactics.

Although these studies look at phonotactic probability in particular, the neighbourhood densities of words are highly correlated with probabilistic phonotactics. Neighbours are defined as words resulting from changing, adding and deleting one of a word's phonemes. A related study by Gathercole (1995) looked at children's performance on a non-word repetition task where non-words were controlled for their relative word-likelihood rating (high or low). Gathercole found that the accuracy of children's productions of non-words with low word-likelihood ratings was related to children's performance on a phonological memory task. For non-words with high word-likelihood rating, children's performance was best predicted by children's lexical knowledge and the size of the children's lexicons. However, the effects seen in Gathercole's experiment might be accounted for by examining the phonotactic probabilities of the non-word stimuli used.

6.1.3. Goal of experiment

Given this body of research, it was asked whether this sensitivity to phonotactic probability would be mirrored in young children's coda productions. To this extent, an experiment was designed that controlled for the frequency of phonemes in nonsense words and that contained the identical coda. The experiment allowed for comparisons of children's productions of the same coda consonant in low and high phonotactic probability environments. The rationale was that if phonotactic probability would differentially affect children's production of the same coda consonant, this would further suggest that the frequency of sounds in the ambient language plays a substantial role in children's acquisition of phonological structures. The hypothesis was that children's production of the same coda consonant would differ according to the environment in which the coda occurred. That is, the same coda will be produced more often in high phonotactic probability non-words than in low phonotactic probability non-words.

6.2. Method

6.2.1. Subjects

Subjects were recruited through the Language Acquisition Lab at the University of Arizona, and were 29 children between the ages of 20–28 months (M=24.76 months), 16 males and 13 females, with a mean MLU of 1.73. All children were monolingual speakers of English with no history of speech or hearing impairment as determined by parent questionnaire (3 of these subjects were exposed to Spanish for no more than 10 hours a week). As described in Chapter 5, parents also completed the MacArthur Communicative Development Inventory (CDI) (Fenson et al, 1984) and subjects had to achieve above the 10^{th} percentile on the Part A: Vocabulary Checklist for inclusion in the experiment (range 10–95^{th} percentile, M=52.4).

Given these criteria, 4 subjects who were tested were excluded from the analyses: 2 subjects were bilingual English and Spanish learners, 1 subject had a history of speech impairment and 1 subject failed to reach the 10^{th} percentile on the Vocabulary portion of the CDI. In addition, 2 subjects were tested, but failed to complete the experiment because of a lack of interest in participating. A subset of the participants also completed the experiment described in the previous chapter.

As with the previous experiment described in Chapter 5, the rationale for including subjects just below and above 2 years of age is that children at this age both produce and delete codas. An attempt was made to include subjects as young as possible (20 months) to determine whether their performance would differ from older children's performance (27–28 months). In fact, younger children performed quantitatively different, but not qualitatively different from older subjects. That is, the same pattern of responses held across the different aged subjects, though the younger participants gave fewer responses overall.

6.2.2. Stimuli

The stimuli consisted of 22 novel CVC words, for which phonotactic probability (low and high), vowel quality (lax, tense and diphthong), and coda type (d, l, s, g, v, ŋ, m) were controlled. Each of the 22 words had a corresponding picture stimulus that consisted of an imaginary animal. The words were presented in a randomly ordered list along with 20 filler nonsense words consisting of VC words. In total there were 11 pairs of low and high phonotactic probability non-words, each with the same vowel length and identical coda. The pairs of words are given in Table 1, The experiment was a within-subjects design and all subjects were tested on the same stimuli.

6.2.2.1. Calculating phonotactic probabilities

The phonotactic probabilities of the non-word stimuli were based on a corpus of CVC words from the CDSC (described in Chapter 3). The CVC word corpus consisted of 604 word types and 40, 822 word tokens. CVC words were marked for whether they were content or function words (e.g., "sock" versus "that"), and whether they were monomorphemic or bimorphemic (e.g., "sock" versus "boys"). This allowed phonotactic probability calculations to be controlled for word class and morphology for the following reasons. One argument against a frequency-based account of phonological acquisition is the fact that /ð/ is frequent in English, yet acquired late (e.g., Moskowitz, 1970, p. 429). However, /ð/ is frequent largely because it appears in English function words (for instance, "the" appears 4,813 times in the CDSC), but children initially delete function words (Brown, 1973). This might account for some of the disparities between frequency and acquisition (also see Bernstein-Ratner, 1994 for a discussion). In addition, research has shown that children differentially produce phonemes depending on their morphological status (Brown, 1973). Thus, words in the corpus were marked for word class and morphological complexity in attempt to control for these factors.

Segmental positional probabilities and biphoneme probabilities were calculated on the log frequency weighted counts (or the token word counts) of words in the CDSC using: (a) all words in the corpus, (b) excluding function words, (c) excluding bimorphemic words, and (d) excluding function words and bimorphemic words. The final lists of C-V-C and CV-VC probabilities were taken from the average segmental positional probabilities and the average biphoneme probabilities across the four ways phonotactic probabilities were calculated (see Table 1). Low and high phonotactic probabilities were defined as segments and biphonemes below and above the median, respectively. Although /tɛs/ "Tess," /dʒɔɪs/ "Joyce," /lɔɪn/ "loin" and /raɪn/ "Rhine" are real words, it was assumed that they are of such low frequency they could function as non-words for the purposes of this experiment. Also, in some dialects of English, [pʌl] is an accepted pronunciation of "pull" and [gɛl] is acceptable for "girl."

Table 1 Stimuli for phonotactic probability (PP) experiment

Coda	Vowel Quality	Low PP	High PP
d	lax	tʃʌd	gɛd
l	lax	pʌl	gɛl
s	lax	dʒʌs	tɛs
g	lax	θæg	sɪg
v	lax	zɛv	dɪv
ŋ	lax	ʃʌŋ	bɪŋ
m	tense	gim	bom
n	tense	von	nin
s	diphthong	dʒɔɪs	faɪs
n	diphthong	lɔɪn	raɪn
d	diphthong	mɔɪd	naɪd

The average segmental positional probability of the low phonotactic probability non-words was .1627 and the high non-words was .1818. The average biphoneme probability of the low phonotactic probability non-words was .0026 and the average biphoneme probability of the high phonotactic probability non-words was .0153. Average probabilities are summarised in Table 2.

Table 2 Average segmental and biphoneme positional phonotactic probabilities for stimuli

Phonotactic probability of non-words	Average Segmental Positional Probabilities	Average Biphoneme Positional Probabilities
High PP	.1818	.0153
Low PP	.1627	.0026

There were a number of assumptions made in the creation of the stimuli. To begin, phonotactic probabilities of CVC words were calculated based only on the CVC words in the input, rather than across words with different syllable structures (CCVC, VC, CVCVC, CVCVC, etc.). Moreover, phonotactic probabilities were calculated on the phonetic transcription of CVC words in isolation, rather than based on phonetic transcriptions taken from continuous speech. In the latter case, effects from coarticulation in running speech have not been coded. Thus, the assumption is that children have access to or representation of words in isolation. Phonotactic probabilities were also calculated at the level of individual phonemes, rather than at more abstract level, such as the featural level. Lastly, phonotactic probabilities were calculated based solely on the adult input. An interesting research question is whether children's own productions mirror phonotactic probabilities in the input, and whether children receive feedback from their own productions.

6.2.2.2. *Neighbourhood density of stimuli*

Neighbourhood densities of the non-word stimuli were also calculated based on the CDSC and the Webster's dictionary. Low phonotactic probability non-words had fewer neighbours than their paired high phonotactic probability non-words; these results are summarised in Table 3.

Table 3 Neighbours of high and low phonotactic probability (PP) non-words

PP of non-words	Average neighbours in CDSC	Average neighbours in Webster's
High PP	11.09	37.73
Low PP	3.8	19.55

Based on the CDSC there were an average of 3.82 neighbouring words (range 0– 7) in the low phonotactic probability non-words, and an average of 11.09 neighbours (range 8–16) in the high phonotactic probability non-words. Similarly, there were fewer neighbours, based on an corpus of adult English words (Webster's), for the low phonotactic probability non-words 19.55 (range 3–27) than the high phonotactic probability non-words, 37.73 (range 9–34). This pattern held for the pairs of words, with one exception based on the neighbours in Webster's, where the low phonotactic probability non-word /pʌl/ had 27 neighbours and the high phonotactic probability non-word /gɛl/ had 23 neighbours.

6.2.2.3. *Word-likelihood ratings of stimuli*

A pilot study was conducted with these stimuli to test whether adults had intuitions about the relative wellformedness of the nonsense CVC words, based on the non-words' probabilities. Fifteen subjects rated the nonsense words on a scale of 1–5 on the words' relative wellformedness or how much they sounded like possible English words (1=relatively poor English word and 5=relatively good English word). The alpha level was set at .05 for all statistical tests. An analysis by subjects found that subjects were more likely to judge high phonotactic probability words as more word-like (M=37.27, SD=8.22) than the phonotactic probability non-words (M=34.4, SD=5.36). This difference was significant, $t(14)$ −1.92, $p<.05$, one-tailed. An items analysis also found that high phonotactic probability non-words were considered as more word-like (M=32.64, SD=4.13) than the low phonotactic probability non-words (M=29.64, SD=8.1). This difference, however, was not significant, $t(10)$ −1.18, p=.13, one-tailed. It is possible that the lack of significant in the items analysis is because some of the non-words stimuli are actual English words, e.g., "Tess" and "loin."[15]

6.2.3. Procedure

The experiment consisted of a repetition task. Children were presented with non-words and their corresponding pictures on a computer using PowerPoint. The stimuli were pre-recorded. Subjects were tested in the Language Acquisition Lab at the University of Arizona and 5 children were tested in their homes. Subjects at the lab were videotaped and subjects tested at home were audio-recorded using a DAT recorder. All sessions were transcribed on-line and checked by a research assistant naïve to the purpose of the experiment. Reliability between the two transcribers was determined by a third person also naïve to the purpose of the experiment and calculated as 98.94% accurate (reliability was based on 17% of the data).

Subjects were told that they would be shown pictures of funny animals and that their job was to repeat the names of the animals. Subjects were given "bear" and "pig" as two practice items. Once the child mastered the task, the experiment began. Subjects were allowed to hear repeated presentations of the stimuli; however, if the child had trouble responding or appeared uninterested, the experimenter would proceed to the next item. At times, children would not repeat the recorded stimulus. In these cases, the experimenter would produce the words for the child to repeat. There was no interaction between the phonotactic probability of the stimuli and the number of times the child heard the stimuli before repeating the word. In addition, there were no differences in children's production of codas in high and low phonotactic probability non-words based on whether the stimuli were presented on the computer versus repeated by the experimenter.

The prediction was that the same coda would be produced more accurately in a high phonotactic probability non-word than in a low phonotactic probability non-word. For example, the prediction was that children would preserve the coda /d/ more in /gɛd/ and would be more likely to delete or modify the coda in /tʃʌd/ (3). A pilot study was conducted with 5 adult subjects to ensure that adults would correctly produce codas in both the low and high phonotactic probability non-words. Of the 5 adults tested, only one coda error on these stimuli was made, in which 1 subject produced /nin/ as [nid].

(3) High PP non-words Low PP non-words

[15] 8 non-native speakers of English were also tested to determine their word-likelihood ratings. There was no significant effect of phonotactic probabilities in both analyses by subjects and by items. An analysis by subjects found that subjects judged high phonotactic probability words (M=35.88, SD=6.01) and low phonotactic probability non-words (M=35.16, SD=5.46) relatively equally. The result was not significant, $t(7)$ −.43, p=.34, one-tailed. Similarly, an items analysis also found that high phonotactic probability non-words were considered equally word-like (M=26.09, SD=5.45) to the low phonotactic probability non-words (M=25.55, SD=6.62). This was also not significant, $t(10)$ −.21, p=.42, one-tailed.

gɛd	tʃʌd
↓	↓
gɛd	tʃʌ, tʃʌt

6.3. Results

6.3.1. Data coding

Responses were coded into five bins, based on broad transcriptions of children's productions: coda produced correctly (Correct), coda produced incorrectly (Incorrect), no coda produced (No Coda), no response (No Response), and real word response (Real Word). The criterion for "real word response" was based on whether the word occurred in the CDSC. Examples of coding are illustrated with the pair /tʃʌd/ and /gɛd/. The responses [tʃʌd] or [tʌd], and [ ] were coded as correct; [tʃʌt], [tʌts], [gɛd] and [gɛdʒ] as incorrect, [tʃʌ] and [g] were coded as No coda responses, and [tʌtʃ] "touch" and [dɛd] "dead" were coded as real word responses. Acoustic measurements and error analyses of children's responses were not done at this time. The results are given by items in Table 4.

Table 4 Responses of correctly produced codas in low and high phonotactic probability (PP) non-words, by items[16]

		Response Type									
		Correct		Incorrect		NoCoda		NR		RealWord	
PP		L	H	L	H	L	H	L	H	L	H
L	H										
tʃʌd	gɛd	7	11	16	8	1	1	2	0	3	9
pʌl	gɛl	3	13	9	2	13	4	0	1	4	9
dʒʌs	tɛs	13	16	4	4	0	0	10	8	2	1
θæg	sɪg	7	14	20	9	1	2	0	0	1	4
zɛv	dɪv	5	5	14	11	2	0	8	10	0	3
ʃʌŋ	bɪŋ	6	11	10	9	3	0	8	6	2	3
gim	bom	8	3	7	6	2	1	1	1	11	18
von	nin	6	14	7	1	3	1	0	0	13	13
dʒɔɪs	fɑɪs	17	18	4	4	2	1	1	3	5	3
lɔɪn	rɑɪn	5	8	3	2	1	4	9	10	11	5
mɔɪd	nɑɪd	8	9	7	7	1	3	12	9	1	1
Total		85	122	101	63	29	17	51	48	53	69

H=High
L=Low

6.3.2. Analyses

Recall that the prediction was that children would be more accurate at producing codas in high phonotactic probability non-words than in low phonotactic probability non-words. The first analyses examined children's accurate coda productions. An analysis by subjects revealed that children were more likely to produce codas accurately in high phonotactic probability words (M=4.21, SD=1.95) than in low phonotactic probability non-words (M=2.93, SD=1.6). Twenty-one children produced more codas in high phonotactic probability words, 5 children produced an equal number of codas in both phonotactic environments, and 3 children produced more codas in the low phonotactic probability words. This difference was significant, *t* (28) −3.71, *p*<.001, one-tailed. Nine of the 11 items were also produced more accurately in the high probability non-words (M=11.09, SD=4.57) than the low probability version (M=7.73, SD=3.98). This difference was also significant, *t*(28) −2.67, *p*<.05, one-tailed.[17]

6.4. Discussion

Children were significantly more likely to produce the same coda in high phonotactic probability non-words than in low phonotactic-probability non-words. These results are consistent with the hypothesis that phonotactic probability is a predictor of coda production in English. Moreover, this finding provides further evidence for the role of the input and the distribution of sound patterns in the ambient language as a basis for phonological acquisition. The results show that children do not simply go from deleting codas to eventually producing codas or even that children will begin to produce more frequent codas before infrequent codas, but importantly, the results show that children will differentially produce the same coda depending on the phonotactic environment.

The results obtained in this experiment are consistent with previous research that has shown that infants, children, and adults are sensitive to these patterns and that they show a preference for non-words composed of high phonotactic probabilities over non-words composed of low-phonotactic probabilities. Previous research has used a variety of tasks such as the Headturn Preference Procedure, judgements of the relative wellformedness of non-words, reaction time in the production of non-words, and a memory task; however, the results here extend this research to the productions of young children between the ages of 20–28 months. The database on which phonotactic-probability was calculated here is unique in that it consisted of adult speech to children (rather than infants) in naturalistic settings (rather than including adult speech from

---

[16] An attempt was made to control for the quality of the onset across the pairs. For example, words were controlled to have stops in onset position across the pairs of low and high phonotactic probability non-words. Although some pairs have different place of articulation (POA) in onset position, children rarely made POA errors in their coda productions.

[17] Children's performance was also coded for their accuracy in producing the word and the rhyme. An analysis by subjects revealed that subjects were more likely to produce rhymes accurately in high phonotactic probability words (M=3.38, SD=1.78) than in low phonotactic probability non-words (M=2.07, SD=1.41). Twenty children produced rhymes more accurately in high phonotactic probability words, six children produced rhymes equally correct across the pairs, and two children produced more rhymes accurately in the low phonotactic probability words. This difference was significant (*t*(28) 3.98, *p*<.001, one-tailed). Eight of the 11 items were produced more accurately in the high probability (M=8.91, SD=3.83) than the low probability version (M=5.45, SD=2.34), which was a significant difference (*t*(10) 2. 91, *p*<.01, one-tailed). At the word level, children were also more accurate at producing high phonotactic probability non-words, than low phonotactic probability words. This was significant both by subjects (*t*(29) −5.97, *p*<.001, one-tailed; and by items: *t*(10) −4.83, *p*<.001, one-tailed).

controlled settings); moreover, probabilities were controlled for word class and morphology (mono- versus bi-morphemic words).

Results differ from those of found by Beckman and Edwards (2000). Their study examined children's productions of word-initial onsets, medial codas and medial clusters, a portion of their stimuli are given in Table 5 below. Non-words were controlled for familiar sequences (occurring often in a corpus of children's spontaneous speech) or novel sequences (legal sequences in English that did not occur in this corpus).

Table 5 Beckman and Edwards stimuli, taken from Beckman and Edwards, 2000, p. 213.

Target	Familiar	Novel [infrequent]
CV	g ˈutən	g ˈautən
VC	gəd ˈotup	gəd ˈotaup
CC	m ˈoftin	m ˈofkən

Beckman and Edwards found effects of Familiar and Novel in children's production of non-words containing CVs and CCs, but not a significant effect for children's production of rhymes, VCs.

The results obtained in this study along with those found in the previous experiment, illustrate the importance of considering the input in young children's acquisition of phonology. Research in this domain often collapses across word types. For example, evidence of the acquisition of /k/ might be taken from children's production of the words "cake" and "toque," even though the frequency of the words "cake" and "toque" and the words' segmental and positional probabilities differ. Consequently, a distorted picture of children's development would be obtained, given that children differentially produce the same consonants depending on the frequency of sound patterns in the whole word.

The findings here require future studies to be more detailed and controlled. When employed, this will allow for a greater understanding of how children acquire language. Specifically, the hope is that this type of research and analysis into frequency will help account for the variability seen in children's acquisition, which has long been argued as an important issue needing explanation. Possibly some of the variation seen in studies results from collapsing words types in the analyses. With finer grained distinctions and analyses of frequency, what was previously unaccounted for as variation might now be seen to reflect the frequency of sounds and words in the ambient language.

Results of the experiment presented here suggest that children's productive language reflects their experience with language more than previously believed in phonological acquisition research. There are a number of ways in which the results can be accounted for, some of which relate to children's processing of language. Recall that the non-words with high probability phonotactics had more neighbours than their matched low probability pair. Children's performance on the non-word repetition task, therefore, might be accounted for by a consideration of the lexicon that the children have already acquired. For example, it might be easier for children to process and retain in memory non-words that were more similar to the words in their lexicon, similar to the work by Gathercole (1995), in which children's production of non-word words with many neighbours was predicted by children's lexicon size. Furthermore, children's accuracy on the non-word repetition task may also reflect children's sublexical representations, as suggested by Beckman and Edwards (2000). Under this account, if children's lexical representations are stored by sub-units, children are then able to access more stable units (based on frequency) with increased accuracy.

CHAPTER 7
Coda acquisition
Discussion and conclusion

7.1.
Introduction

The central goal of this dissertation has been to explore two theories of language acquisition. The first theory argues that language acquisition is primarily mediated by innate properties of language provided by UG. The other theory argues that language acquisition is best described with respect to frequently occurring patterns in the ambient language. The domain in which these theories were contrasted was children's acquisition of coda consonants, specifically, word-final consonants in CVC words. In this chapter, I provide a summary of how these two theories of acquisition were tested and how the conclusions were drawn. I then discuss the methodological implications of the present studies, research implications for theories of acquisition, and future research directions.

7.2.
Review

7.2.1.
Review of the Universal Grammar Hypothesis (UGH)

The starting point of this research was based on the observation that child language reflects cross-linguistic markedness. Using examples drawn from syllable typology, it was illustrated that the least marked syllable structure is a CV syllable, such that languages exist that only have CV syllables. Parallel to this, children initially delete final consonants; thus, their productions conform to CV syllable shapes and reflect the universally preferred syllable shape. The traditional account is that this observation reflects UG.

UG (established primarily through cross-linguistic research) has been argued to constitute the initial hypotheses children entertain for language. Thus, if the universally unmarked syllable shape is a CV syllable, children should exhibit an initial sensitivity to these syllable types. To explore a UG-based account of coda consonant acquisition it was necessary to establish what is universally preferred in coda position. These patterns were based on previously published research and based on frequency analyses of word-final consonants. Results showed that languages prefer codas with coronal place of articulation and increased sonority (Chapter 3). These cross-linguistic preferences were interpreted as reflecting the universally preferred codas or UG.

Thus, the prediction of coda consonant acquisition in English based on UG was formulated as:

(1) Universal Grammar Hypothesis (UGH_2): Children will produce coronal codas (t, d, s, z, n, l, r) more than labial or dorsal codas, and sonorant codas (m, n, ŋ, l, r) more than obstruent codas in word-final position.

Based on previous parallels drawn between cross-linguistic markedness and child language, children should show a sensitivity to what is unmarked in coda position.

7.2.2. Review of the Specific Language Grammar Hypothesis (SLGH)

It was noted that one must consider the fact that cross-linguistic markedness is generally mirrored in a single language's distribution of sounds. This brings into question the relationship between child language and UG; child language patterns can also be characterised as reflecting the distribution of sounds in the ambient language. To explore whether the input could account for patterns seen in children's coda consonant acquisition, an examination of English codas was provided in Chapter 3. An argument was made that an input-based account of language acquisition and coda consonant acquisition should be based on the most frequent codas in the ambient language:

(2) Specific Language Grammar Hypothesis (SLGH): Children will produce the most frequent codas in word-final position: t>r>n>d> z>k>s>l>m>v>ʃ>g>p>θ>ŋ>tʃ>f>dʒ>b>ʒ, ð.

To test which of these two hypotheses better predicted children's acquisition of coda consonants, data were collected from children acquiring English.

7.2.3. Review of child language data

Acquisition data from previously published research and from existing data from CHILDES were analysed to establish the first codas children produce (Independent Analysis) and the first codas children produce correctly (Relational Analysis) (Chapter 4). In addition, an experiment was designed to test children's production of a variety of coda consonants in English (Chapter 5). The child language patterns from the different sources are summarised in Table 1.

Table 1 Child Language patterns English (Independent Analysis, Relational Analysis and Experiment 1)

Analysis	Child Language Patterns: English
Independent Analysis	t>n>k>d, m>ʔ>p>tʃ, l> b, f, F, ŋ, r, s
Relational Analysis	k>t>n>d, s>b, f, l, m, ŋ, p, r, z
Experiment 1	r>ʃ>t, n, k>d>p>m>θ>b>ŋ, g

It was then possible to evaluate the UGH and the SLGH with respect to the patterns seen in child language.

7.2.4.
Results and discussion

Children's productions of codas and their relationship to the two hypotheses are summarised in Table 2 below. "✘" represent no relationship between the predicted and attested patterns in acquisition. "✓" indicates that the predicted patterns were mirrored in children's production of coda consonants.

Table 2 Children's coda productions in English from different sources and how they relate to the UGH and SLGH

Analysis	UGH	SLGH
Independent Analysis	✘	✓
Relational Analysis	✘	✓
Experiment 1	✘	✓

7.2.4.1.
Results and discussion: UGH

To evaluate child language with respect to the UGH, children's coda productions were analysed to determine whether the preferred codas had coronal place of articulation and whether they were sonorant codas. Results showed no preference for these codas in children's productions, thereby, refuting the UGH.

It is an open question, however, whether other interpretations of UG can capture these data. Take, for example, the observation from §2.6 that any UGH would not realistically predict that children should first produce the segments /tʃ/ and /dʒ/, despite the fact that they bear coronal place of articulation. This is because there are many different dimensions that can determine which segments are marked and unmarked; for example, /tʃ/ and /dʒ/ are also affricates, which makes them more marked segments in terms of their manner. It should be possible to develop a UGH account of coda acquisition taking into consideration the various features that segments can bear and the markedness of features, such as voice, continuancy, sonority, and place of articulation. The best measure for evaluating or determining which features or which feature combinations are marked, however, is not clear. In implementing such a measure, one needs criteria for determining what is marked, but the criteria for markedness can vary across languages. For example, a segment may be more or less marked with respect to the other segments in a language's segmental inventory. In addition, to develop an order of acquisition based on a measure of this type, one also needs to consider other factors such as positional markedness, where the markedness of features and feature combinations vary according to segments' position within prosodic structure. For example, a segment may be marked in one position, but not in another position (e.g., coda inventories vary according to syllable stress and position within a word). Given its complexity and size, this question is left to future research. However, it is important to note that the most common conception of UGH is not adequate in explaining children's early coda productions, as this dissertation shows.

7.2.4.2.
Results and discussion: SLGH

To evaluate children's coda productions against the SLGH, analyses determined whether there were significant correlations between the codas produced by children and the frequency of those codas in English. Across all three operalisations of child language data, there was a significant relationship between children's coda productions and the frequency of codas in English. Thus, as formalised in this dissertation, the results favour the SLGH over the UGH as an account of children's acquisition of coda consonants in CVC words. Based on the evidence provided in these analyses, children's coda acquisition is best characterised in terms of patterns in the input rather than innate UG. In this sense, children organise and build their phonological knowledge of word-final consonants upon frequently occurring patterns in the ambient language. Because children do not show evidence for the universally preferred codas in their productions, this suggests that children do not necessarily come to the acquisition task with prespecified knowledge of what the optimal codas are.

The results also present an interesting finding for the properties of the input upon which children calculate frequency. Recall that in Chapter 3, English was analysed with respect to the universal preferences of coronal place of articulation and sonority in coda position. Results showed that with place of articulation, both EFA and AFA analyses showed evidence for the favoured status of coronal place of articulation (see Appendices VII and VIII). Results of sonority revealed that English has more sonorant codas than expected (EFA), and a greater number of obstruent codas (AFA) (see Appendices IX and X). Interestingly, there are different coda preferences depending on the analysis; there is a larger proportion of sonorant codas than expected, but a significantly greater number of obstruent than sonorant codas in CVC words. As in Chapter 2, this latter finding can be accounted for by appealing to English's coda inventory, which has more obstruents (16) than sonorants (5). It has been argued in this dissertation that children do not show a preference for the unmarked codas that are sonorant, thus refuting the UGH. This further suggests, however, that children do not attend to proportions of obstruents and sonorants in the input (EFA), but rather, that children attend to input frequencies based on the raw number of occurrences of codas in the language (AFA).

Recall that previous research looking at coda consonant acquisition has noted that obstruents are acquired before sonorants in coda position (Fikkert, 1994; Goad, 1997; Salidis & Johnson, 1997) and that voiceless obstruents and nasals are preferred over voiced obstruents in coda position (Bernhardt & Stemberger, 1998; Kehoe & Stoel-Gammon, 2001). Accounts for this have argued that children produce the maximal contrast in coda position from the preceding vowel (Fikkert, 1994), that this reflects the structure of these segments in which they are syllabified as the onsets of empty headed syllables (Goad, 1997), or that this reflects the default coda (margin) features and articulatory and perceptual ease (Bernhardt & Stemberger, 1998; Kehoe & Stoel-Gammon, 2001). A closer look at the patterns from previous research (and noted in previous research) reveals that children's production of codas is often varied, and children produce a variety of segments as their initial coda productions. Thus, looking at acquisition in terms of classes of segments is an abstraction from the data which might inhibit the developmental picture. When we look at the words and segments that children are mostly like to acquire and produce as their first words, an additional account to the patterns can be found. Here the patterns in the ambient language drive the acquisition effects.

To further examine the role of the input in children's acquisition of coda consonants, a study was designed to test children's production of the identical coda in non-words which were controlled for their phonotactic probabilities based on a corpus of child-directed speech (Chapter 6). Results showed that probabilistic phonotactics played a significant role in accounting for children's coda consonant productions and illustrate the importance of considering the role of linguistic input in children's acquisition of phonological structures.

7.2.4.3. Accounting for frequency effects in acquisition

The results present a number of questions about the nature of language acquisition and the assumptions made in child language acquisition research. To begin, this research raises questions about the mechanisms involved in language acquisition. The UGH as formulated in this dissertation was refuted, and instead, the conclusion drawn is that the acquisition of coda consonants is best characterised with respect to frequently occurring sounds in the input. The question remains as to what mechanisms can then account for children's sensitivity to frequently occurring patterns. Any theory of acquisition which relies on the input to organise children's linguistic representations needs to ultimately account for how and why frequently occurring properties are attended to by children. Demonstrating that children are sensitive to frequency information does not explain how children achieve this. In this sense, input-based models of acquisition are plagued with the same problems found in UG-and innate-based accounts of acquisition.

To address this question, there are a number of methods that might begin to provide answers. Research in cognitive development other than language development, and research comparing human development to that of other primates might help address what aspects of development are specific to language, and specific to humans. Moreover, studies such as the ones presented in this dissertation, which provide a detailed exploration of the patterns in the input and children's sensitivity to these patterns, can further describe children's abilities and how these might change over time. These data can also be potentially captured using connectionist models, which might further provide insight into the acquisition task.

Children's ability to produce particular codas better than others can be characterised in a number of ways. This ability might reflect children's experience with aspects of language, such as their acquired lexicon. In this case, children may have more detailed lexical representations of frequent words (and subsequently, the consonants). Similarly, when it comes to production, these words' representations may have an advantage in that they can be accessed faster and more accurately than other words that might not have as detailed representations. By establishing the aspects of the input that children attend to, the hope is also that this research will further elucidate how children represent and process language. For example, future research could further explain the units that are used in language representation and lexical access, and account for how language is perceived and produced through the mapping of acoustics to articulatory motor abilities.

7.2.4.4. Alternative account: Ease of perception and/or production

One possibility is that the findings presented here reflect a universal "ease of articulation" or "ease of perception" in coda position, rather than the frequency of codas in the input language. One of the largest difficulties in evaluating children's coda productions with respect to these alternative hypotheses, is that there is no scale or definition for ease of articulation or perception. In addition "ease" can differ according to experience, such that frequent sounds that children have experience producing and perceiving might present less difficulty than infrequent sounds (Bernhardt & Stemberger, 1998), thus, accentuating the potential importance of frequency explanations.

There are a number of reasons that suggest that articulatory and perceptual ease cannot account for the results described in this dissertation. To begin, a similar effect is found before children begin to speak (Jusczyk et al., 1994), in that infants at 9 months of age show a sensitivity to nonsense words composed of more frequent phonemes and phoneme combinations than nonsense words that have infrequent sounds.

Thus, at a stage where articulatory effects are not applicable, the same sensitivity to frequency in the input language is found.

Another fact suggesting that articulatory or perceptual constraints cannot explain the patterns seen in coda consonant acquisition comes from the different patterns seen in children's spontaneous versus imitated production of codas. A motor or auditory prediction of children's responses cannot account for why these patterns are distinct. Similar paradoxes for articulatory-based theories of acquisition come from research that shows children are more accurate and phonologically advanced in producing sounds contained within words than are unanalysed forms, such as nursery rhymes or songs.

Additional evidence against both articulatory ease and perceptual ease accounts of acquisition comes from cross-linguistic acquisition research illustrating that children who are learning different languages acquire phonemes at different rates depending on the frequency of the sounds in the ambient languages (as seen in §1.3.1). For example, children learning Spanish acquire /tʃ/ earlier than children learning English, arguably because /tʃ/ is more frequent in Spanish than in English (Macken, 1995). Similar arguments have been made for the acquisition of a variety of phonemes across languages. In addition, evidence was reviewed where children learning English acquire codas before children learning Spanish. Again, this can be accounted for with respect to the frequency of codas in English versus Spanish, where English has a higher frequency of codas. One caveat, however, is that observations of phoneme acquisition in different languages are often based on naturalistic data. Thus, these results may reflect the sounds of the language that children are most likely to attempt. In other words, children learning English may not attempt many words containing /tʃ/, which might account for why the sound does not appear in children's productions. If one were to test children's production of different phonemes or codas in both languages, one may find that children's performance increases. These methods could then be employed in cross-linguistic acquisition data to further distinguish between articulatory and perceptual ease effects in acquisition.

7.3. Future directions

The patterns seen here in children's production of coda consonants raise a number of interesting questions about the nature of children's lexical representations and the basis on which children acquire their ambient language's phonology. Future research is needed in order to further distinguish and qualify this development. The following discussion provides some logical places from which research can begin.

7.3.1. Prosodic and segmental interactions, and word class

It has been mentioned a number of times throughout this dissertation that there are segmental and prosodic interactions affecting the patterns of codas in English. For example, coda distributions differ according to whether codas appear at the end of stressed or unstressed syllables (Zamuner, 1998a, 1998b). An examination of children's sensitivity to frequency and phonotactic probabilities in these different environments might further reveal whether children keep these environments distinct or whether children average frequencies across them.

Other extensions of this research could be applied to word class. For example, all the words in the experiment in Chapter 5 were restricted to CVC content words. An immediate extension would test children's coda productions across word categories. Research has found differential phoneme distributions and stress patterns in words with different classes (e.g., in nouns and verbs), and in adults and children's sensitivities to these patterns (e.g., Camarata & Leonard, 1985, 1986; Camarata & Schwartz, 1985; Cassidy & Kelly, 1991; Kelly, 1996, Sereno & Jongman, 1990; Sherman, 1975). If infants and children are tested on their sensitivity to phoneme distributions that are controlled for word class, this might reveal whether they encode this information and whether they can use this information in acquiring language. For example, if children are sensitive to the frequency of different sound patterns in nouns versus verbs, when children encounter a novel word they might be able to hypothesise the word's class based on the frequency and distribution of the word's sounds.

7.3.2.
Optimality Theory

A current approach to modelling children's acquisition of phonological structures is within Optimality Theory (OT) (McCarthy & Prince, 1993; Prince & Smolensky, 1993). OT is a phonological theory that consists of both an input and an output. The relationship between these representations is determined by universal violable constraints, whose ranking is determined on a language specific basis. In the model, the set of constraints (CON) is given by UG and an infinite number of potential candidates are provided by a generator (GEN). The input and output candidates are then filtered through these constraints. The evaluator (EVAL) selects the optimal output based on the criteria of constraint filtering. Because the set of constraints is part of every languages' grammar, constraints will be both conflicting and violated in outputs; even the optimal candidate will violate lower ranked constraints.

Important to the theory is the notion that constraints are innately provided by UG, and that cross-linguistic variation arises from constraint rankings. Also central to the theory is the two major classes of constraints: wellformedness and faithfulness constraints. Wellformedness constraints require outputs to conform to certain specifications, whereas faithfulness constraints restrict outputs to be true to their input specification.

As a theory of phonological acquisition, OT appeals to innate constraints, and different characterisations of OT focus on the initial ranking of these constraints, and how constraints are reorganised as the child acquires the adult grammar. For example, within the theory, researchers debate whether wellformedness or markedness constraints are ranked above faithfulness constraints (Bernhardt & Stemberger, 1998; Curtin, in press; Demuth, 1995; Goad, 1997b, Gnanadesikan, 1995; Ohala, 1994, 1996; Smolensky, 1996), whether the class of faithfulness constraints initially outranks markedness constraints (Hale & Reiss, 1995), or whether there is no initial constraint ranking (Tesar 1995; Tesar & Smolenksy 1993).

The initial characterisation of acquisition within OT was one in which wellformedness or markedness constraints were initially ranked above faithfulness constraints (Gnanadesikan, 1995, Ohala, 1994, 1996). The observation that child language resembles cross-linguistic variation is captured by markedness constraints ranked initially high; thus, children's productions tend to be unmarked (also termed "emergence of the unmarked"). For example, Ohala (1996) characterises onset cluster reduction in child language as the product of a markedness constraint against complexity, *COMPLEX: Syllables must not associate more than one C or V to a syllable position node. Thus, children's cluster reductions is formalised as the result of the innate constraint *COMPLEX, which forces children to conform onset clusters to singleton consonants (3).

(3)	Child	Adult		
	[kin]	/klin/	*clean*	(Gnanadesikan, 1995)
	[piz]	/pliz/	*please*	(Gnanadesikan, 1995)

[tig]	/stig/	(Ohala, 1996)
[fɪs]	/fɪsk/	(Ohala, 1996)

Constraints on sonority then govern which consonant of the cluster is deleted. Acquisition of the adult grammar is a process of promoting faithfulness constraints over markedness constraints such that the child's production will eventually be faithful to the adult target, and the constraint *COMPLEX is low ranked and not evidenced in the output.

Moreover, approaches within OT differ according to how the child acquires the correct constraint ranking for their specific language. For example, the model adopted by Ohala (1994, 1996) interprets the child as having to rank faithfulness constraints above markedness constraints, whereas Tesar & Smolensky (1993, 1998) argue for constraint demotion (error driven constraint demotion algorithm).

In addition, within the OT framework, researchers debate whether the constraints in the grammar are strictly ranked (Tesar & Smolensky, 1993, 1998; Prince & Smolensky, 1993) or stochastic (Boersma, 1998; Hayes & MacEachern 1998). For example, Boersma's model (1997, 1999) allows for variability and optionality, and it can encode into the grammar the relative frequency of variants. Because most of the these approaches within acquisition focus on children's initial word production as the beginning point to model the ordering of constraint ranking (cf. Boersma), they often fail to account for the variability seen in children's productions at the onset of word production (Velleman, 2000). For example, a large body of research illustrates that infants have knowledge of their language's phonotactic patterns at a young age and this suggests that children's prior exposure to language has consequences for constraint ranking in their grammar (see Curtin, 2001).

A strength of OT is its ability to capture the relationship between cross-linguistic markedness and child language. Parallels between adult and child language result from a universal set of constraints and from the initial ranking of constraints in an unmarked order (Goad, 1997b). Common across the different characterisations of OT is their appeal to innate constraints, which are provided by UG (but see Boersma, 1998 for a functional approach to learning constraints). Thus, children are born with innate knowledge of how language is structured and this is the primary motivation for children's modification of the input. In most OT acquisition research, it is assumed that children have adult like perceptual systems and that they have access to the adult output representations. The goal of these theories again, is to capture how the child learns the constraint ranking of the target language.

The question, however, then arises to the utility of OT as a theory of acquisition, given the evidence presented in this dissertation showing that when there is a disparity between what is cross-linguistically unmarked and the distribution of sounds in the input, the input serves to drive children's acquisition of phonological structures. A central premise of OT is that children are born with innate markedness constraints, yet children appear to demonstrate sensitivity to the input over the unmarked properties of codas in their production of coda consonants, and that infants demonstrate sensitivity to these same patterns in infancy (see §7.3.5). These factors suggest that models of OT might be adapted to account for acquisition at a younger age, where markedness might play more of a substantial role in phonological acquisition.

Were these data modelled in OT, they would require a stochastic implementation of the theory, where the frequency of the input would be used to rank innate constraints as children move from an initial state towards the adult grammar (see Boersma & Levelt, 1999 for the acquisition of syllable shape; Curtin & Zuraw, 2001 and Curtin, 2001 for prosodic development). Results from the experiments in this dissertation, however, illustrate that the model must also allow for a way of referring to the probability of specific segments and biphonemes in acquisition. That is, the model must capture the fact that children's production of /d/ in coda position can vary according to the overall phonotactic probabilities of the words that have the coda /d/ (see Hammond, 2001 for an implementation of OT which considers such factors).

7.3.3. Clinical extensions

The results presented here also have implications for clinical work. Because children differentially produce the same coda depending on the word's phonotactic probability, frequency needs to be controlled for in standardised assessments of children's and adults' expressive language and/or consonant inventories. If the materials used to estimate performance do not control for these factors, an inaccurate assessment might be obtained. It would also be beneficial for training methodology to determine whether children with phonological delay, who are treated on target sounds in infrequent environments, can then transfer their improvement post-treatment to the same target sounds in more frequent environments or vice versa. In addition, further research looking at typically developing children between the ages of 20–28 months (with increased numbers of subjects at a younger age) will further solidify the patterns of young children's sensitivity to coda frequency and phonotactic probability in production. This can then provide a baseline from which children with language impairment can be compared.

7.3.4. Logical possibilities for language acquisition

As mentioned in Chapter 1, there are a number of logical possibilities of how the domains of innate UG and the input might interact in children's acquisition of phonology. In Table 3 (repeated from Chapter 1), possible patterns for the interplay between innate UG, the input (Language Specific Patterns), and child language are given. "Same" denotes similar patterns across domains, and "different" signifies a domain having a unique pattern.

Table 3 Logical possibilities for language acquisition

	UG (X-Linguistic data)	Language Specific Patterns (LSP)	Child Language Patterns (CLP)
a.	same	same	same
→b.	same	different	same
→c.	different	same	same
d.	same	same	different
e.	different	different	different

In the case of (a), the same patterns are seen across all three domains, which is not informative with respect to whether child language is the result of innate UG or the input (LSP). The following cases of (b) and (c) were explored in this dissertation because these are cases where child language patterns are potentially disassociated from UG or the input. The examination of place of articulation and sonority showed that these are cases of (b), where the input patterns are somewhat distinct from those of cross-linguistic markedness. In these cases, children pattern with the input over UG. It remains a question for future research whether there is evidence of (c) in acquisition.

The last two patterns of (d) and (e) where children's patterns depart from UG and the input were also not explored in this dissertation. These cases may arise from children's cognitive or physical limitations, where despite some sounds being cross-linguistically unmarked and frequent in languages, children are unable to perceive or produce certain sounds at a young age. As seen in the following section, further research is needed to determine whether children's abilities and sensitivities to UG and/or the input remain constant or change over time.

7.3.5. Phonological development and Experiment 3: Infant study

Another consideration would be to examine children's production of coda consonants to determine whether there are quantitative and/or qualitative differences in development. There were too few data collected here from younger subjects to do an analysis on subjects' performance based on age. It is possible that the UGH could account for children's coda productions at a younger age, or it is also possible that the frequency of codas in the input language are more or less reflected in children's productions across development.

Related to this, in addition to testing 2-year-olds production of the non-word stimuli created, 7-month-old infants were also tested on whether they could differentiate between the low and high phonotactic probability CVC non-words described in Chapter 6. Previous research by Jusczyk et al. (1994) demonstrate that 9-month-old infants, but not 6-month-old infants, are sensitive to phonotactic probabilities on CVC non-words, such that they listen longer to lists of words composed of high phonotactic probabilities. The goal was to determine whether the same pattern would be found with the stimuli created here. Twenty 7-month-old infants were exposed to lists of low and high phonotactic probability CVC words. The Headturn Preference Procedure was used to determine children's preferences for the lists of low and high phonotactic probability non-words. Results showed that infants showed a significant preference for CVC non-words with high phonotactic probabilities (M=9.66, SD=2.95) over those with low phonotactic probabilities (M=8.47, SD=1. 94), $t(19)$ 2.66, $p<.05$, one-tailed. This replicates the finding shown by Jusczyk et al, (1994), yet extends it to infants who are two months younger. This illustrates that even by 7-months-of-age, infants are able to distinguish between frequent and infrequent sounds and sound combinations in the ambient language. This also suggests that children's sensitivity to the input begins before production, such that the patterns of coda consonant development seen in this dissertation reflect the learner's sensitivity to the input at a very young age.

7.4. Conclusion

The research explored in this dissertation supports an input-based account of phonological acquisition. Under this view, language acquisition is best characterised with respect to the patterns in the ambient language, where frequently occurring properties of the input serve to organise children's linguistic representations.

Or perhaps clock oil is responsible for the patterns seen in young children's productions, and in this case, children should just eat fewer clocks!

> In a country the other side of tomorrow, an ogre who had eaten a clock and had fallen into the habit of eating clocks was eating a clock in the clockroom of his castle when his ogress and their ilk knocked down the locked door and shook their hairy heads at him.

"Walsa malla?" gurgled the ogre, for too much clock oil had turned all his "t"s to "l" s.

"Just look at this room!" exclaimed the ogress, and they all looked at the room, the ogre with eyes as fogged as the headlights of an ancient limousine. The stone floor of the room was littered with fragments of dials, oily clocks and springs, broken clock hands, and pieces of pendulums. "I've brought a doctor to look at you," the ogress said.

The doctor wore a black beard, carried a black bag, and gave the ogre a black look. "This case is clearly not in my area," he said.

The ogre struck three, and the doctor flushed.

"This is a case for a clockman," the doctor said, "for the problem is not what clocks have done to the ogre but what the ogre has done to clocks."

"Wulsa malla?" the ogre gurgled again.

"Eating clocks has turned all his "t"s to "l" s," the ogress said. "That's what clocks have done to him."

"Then your clockman may have to call in consultation a semanticist or a dictionist or an etymologist or a syntaxman," the non-clock doctor said, and he bowed stiffly and left the room.

"The Last Clock"

James Thurber (1961)

APPENDIX I
Languages used in cross-linguistic analyses organised by language

Language	Language Family[18]	Reference
Alawa	Australian	Sharpe, 1972
Arabic	Afro-Asiatic	Abu-Absi, 1995
Armenian	Indo-European	Kozintseva, 1998
Brahui	Dravidian	Bray, 1986
Cantonese	Sino-Tibetan	Huang, 1970
Daga	Trans-New Guinea	Murane, 1974
Daur	Altaic	Wu, 1998
Dehong	Daic	Luo, 1998
Dehu	Austronesian	Tryon, 1967a, 1967b
Doyayo	Niger-Congo	Wiering & Wiering, 1994
Dutch	Indo-European	Baayen, Piepenbrock & Rijn, 1993
Hebrew	Afro-Asiatic	Doniach & Kahane, 1996
Jacaltec	Mayan	Craig, 1977
Karachay	Altaic	Seegmiller, 1996
Khmer	Austro-Asiatic	Headley, Chhor, Lim, Cheang & Chun, 1977
Kurdish	Indo-European	McCarus, 1967

cont.

Language	Language Family	Reference
Latvian	Indo-European	Fennell, & Gelsen, 1980; Nau, 1998; Turkina, 1982
Lisu	Sino-Tibetan	Bradley, 1994
Louisiana Creole French	Creole	Valdman, Klinger, Marshell & Roltet, 1988
Luo	Nilo-Saharan	Stafford, 1967
Mam	Mayan	England, 1983
Mbay	Nilo-Saharan	Keegan, 1997; Keegan, Marcel & Bertin, 1996
Munari	Austro-Asiatic	Sinha, 1975

[18] Language family classifications taken from Ethnology (Grimes & Grimes, 1999).

Language	Language Family	Reference
Nivkh (Gilyak)	Language Isolate	Gruzdeva, 1998
Pima Bajo	Uto-Aztecan	Estrada Fernández, 1996
Pulaar	Niger-Congo	Niang, 1997
Rhade	Austronesian	Tharp & Y-Bham, 1980
Salt-Yui	Trans-New Guinea	Irwin, 1974
Serbian	Indo-European	Benson, 1982
Telefol	Trans-New Guinea	Healey, 1977
		Healey & Healey, 1977
Thai	Daic	Manich Jumsai, 1958
Tibetan	Sino-Tibetan	Goldstein & Nornang, 1970
Totonac	Totonacan	MacKay, 1999
Turkmen	Altaic	Garrett, Lastowka, Naahielua & Pallipamu, 1996
Ukrainian	Indo-European	Danylenko & Vakulenko, 1994

APPENDIX II
Languages used in cross-linguistic analyses organised by language family

Language Family[19]	Language
Afro-Asiatic	Arabic
Afro-Asiatic	Hebrew
Altaic	Daur
Altaic	Karachay
Altaic	Turkmen
Australian	Alawa
Austro-Asiatic	Khmer
Austro-Asiatic	Munari
Austronesian	Dehu
Austronesian	Rhade
Creole	Louisiana Creole French
Daic	Dehong
Daic	Thai
Dravidian	Brahui
Indo-European	Armenian
Indo-European	Dutch
Indo-European	Kurdish
Indo-European	Latvian
Indo-European	Serbian
Indo-European	Ukrainian
Language Isolate	Nivkh (Gilyak)
Mayan	Jacaltec
Mayan	Mam
Niger-Congo	Doyayo
Niger-Congo	Pulaar
Nilo-Saharan	Luo
Nilo-Saharan	Mbay

[19]Language family classifications taken from Ethnology (Grimes & Grimes, 1999).

Language Family	Language
Sino-Tibetan	Cantonese
Sino-Tibetan	Lisu
Sino-Tibetan	Tibetan
Totonacan	Totonac
Trans-New Guinea	Daga
Trans-New Guinea	Salt-Yui
Trans-New Guinea	Telefol
Uto-Aztecan	Pima Bajo

APPENDIX III
EFA for cross-linguistic place of articulation in coda position

A. Languages with more coronal codas than expected
*$p<.1$, **$p<.05$, *** $p<.01$, **** $p<.001$

Alawa*
Arabic*
Armenian*
Brahui****
Cantonese
Daga
Daur****
Dehong*
Doyayo****
Dutch****
Hebrew
Jacaltec***
Karachay*
Kurdish****
Latvian****
Lisu****
Louisana French Creole****
Mbay***
Nivkh***
Pima Bajo
Pulaar**
Salt-Yui***
Telefol
Tibetan***
Thai
Totonac***
Turkmen**

B. Languages with fewer coronal codas than expected
*$p<.1$, **$p<.05$, ***$p<.01$, ****$p<.001$

Dehu
Khmer****
Luo
Mam
Munari
Rhade****
Serbian***
Ukranian**

APPENDIX IV
AFA for cross-linguistic place of articulation in coda position

A. Languages with more coronal codas than labial and dorsal codas

Alawa	Louisana French Creole
Arabic	Luo
Armenian	Mam
Brahui	Mbay
Daur	Munari
Dehu	Nivkh
Doyayo	Pima Bajo
Dutch	Pulaar
Hebrew	Salt-Yui
Jacaltec	Serbian
Karachay	Totonac
Kurdish	Turkmen
Latvian	Ukranian
Lisu	

B. Languages with fewer coronal codas than labial and dorsal codas

Cantonese	Rhade
Dehong	Telefol
Daga	Thai
Khmer	Tibetan

APPENDIX V
EFA for cross-linguistic sonority in coda position

A. Languages with more sonorant codas than expected
*p<.1, **p<.05, ***p<.01, ****p<.001

Arabic****
Armenian
Brahui
Cantonese***
Daga
Dehong**
Dehu****
Doyayo
Jacaltec
Khmer****
Kurdish*
Louisana French Creole**
Mam
Nivkh
Pulaar***
Rhade*
Telefol
Thai***
Tibetan****
Totonac****
Turkmen****
Ukranian

B. Languages with fewer sonorant codas than expected
*p<.1, **p<.05, ***p<.01, ****p<.001

Alawa
Daur
Dutch**
Hebrew
Karachay
Latvian*
Lisu****
Luo*
Munari
Serbian

C. Languages with number of sonorant codas as expected

Mbay
Pima Bajo
Salt-Yui

APPENDIX VI
AFA for cross-linguistic sonority in coda position

A. Languages with more sonorant codas than obstruent codas

Arabic	Pulaar
Cantonese	Rhade
Dehong	Salt-Yui
Khmer	Telefol
Mbay	Thai
Munari	Tibetan

B. Languages with fewer sonorant codas than obstruent codas

Alawa	Latvian
Armenian	Lisu
Brahui	Louisana French Creole
Daga	Luo
Daur	Mam
Doyayo	Nivkh
Dutch	Pima Bajo
Hebrew	Serbian
Jacaltec	Totonac
Karachay	Turkmen
Kurdish	Ukranian

C. Languages with equal number of sonorant and obstruent codas

Dehu

APPENDIX VII
EFA for coronal place of articulation in English

EFA for coronal place of articulation for different sources in English (see §3.2)

English Source	χ^2 Goodness of Fit results	Expected coronal *f*	Observed coronal *f*
Random House	χ^2 (1, N=2001)=4.17, $p<.05$.	1238.62	1283
Webster's	χ^2 (1, N=1299)=4.94, $p<.05$.	804.08	843
CDI	χ^2 (1, N=152)=.43, $p=.5$.	94.08	98
CDSC-type	χ^2 (1, N=604)=10.75, $p<.01$.	373.88	413
CDSC-token	χ^2 (1, N=40822)=5629.93, $p<.001$.	25268.82	32631

APPENDIX VIII
AFA for coronal place of articulation in English

AFA for coronal place of articulation for different sources in English (see §3.2)

English Source	χ^2 Goodness of Fit results	Observed coronal *f*	Observed labial and dorsal *f*
Random House	$\chi 2$ (1, N=2001)=159.53, *p*<.001.	1283	718
Webster's	$\chi 2$ (1, N=1299)=115.3, *p*<.001.	843	456
CDI	$\chi 2$ (1, N=152)=12.74, *p*<.001.	98	54
CDSC-type	$\chi 2$ (1, N=604)=81.59, *p*<.001.	413	191
CDSC-token	$\chi 2$ (1, N=40,822)=14632.15, *p*<.001.	32631	8191

APPENDIX IX
EFA for sonority in English

EFA for sonority for different sources in English (see §3.2)

English Source	χ^2 Goodness of Fit results	Expected sonorant *f*	Observed sonorant *f*
Random House	χ^2 (1, N=2001)=183.09, p<.001.	476.29	734
Webster's	χ^2(1, N=1299)=64.05, p<.001.	309.16	432
CDI	χ^2 (1, N=152)=1.23, p<.05.	36.18	42
CDSC-type	χ^2 (1, N=604)=29.92, p<.001.	143.75	201
CDSC-token	χ^2 (1, N=40822)=2951.34, p<.001.	9715.64	14390

APPENDIX X
AFA for sonority in English

AFA for sonority for different sources in English (see §3.2)

English Source	χ^2 Goodness of Fit results	Observed sonorant *f*	Observed obstruent *f*
Random House	χ^2(1, N=2001)=141.97, p<.001.	734	1267
Webster's	χ^2 (1, N=1299)=145.67, p<.001.	432	867
CDI	χ^2(1, N=152)=30.42, p<.001.	42	110
CDSC-type	χ^2 (1, N=604)=67.56, p<.001.	201	403
CDSC-token	χ^2 (1, N=40,822)=3552.25, p<.001.	14390	26432

References

Abu-Absi, S. (1995). *Chadian Arabic.* Munich, Germany: Lincom Europa.

Albright, R.W., & Albright, J.B. (1956). The phonology of a two-year-old child. *Word, 12,* 382–390

Baayen, R.H., Piepenbrock, R., & Rijn, H.van. (1993). *The CELEX lexical database* [CD-ROM]. Philadelphia: Linguistics Data Consortium, University of Pennsylvania.

Barrett, M., Harris, M., & Chasin, J. (1991). Early lexical development and maternal speech: a comparison of children's initial and subsequent uses of words. *Journal of Child Language, 18*(1), 21–40.

Bates, E., Bretherton, I., & Snyder, L. (1988). *From first words to grammar: Individual differences and dissociable mechanisms.* Cambridge, MA: Cambridge University Press.

Beckman, J. (1998). *Positional faithfulness.* Doctoral dissertation, University of Massachusetts, Amherst. New York: Garland Press, 1999.

Beckman, M.E., & Edwards, J. (2000). Lexical frequency effects on young childrern's imitative productions. In M.Broe & J.Pierrehumbert (Eds.), *Papers in Laboratory Phonology V: Language Acquisition and the Lexicon* (pp. 208–217). Cambridge: Cambridge University Press.

Benson, M. (1982). *Standard English-SerboCroatian, Serbo-Croatian-English dictionary.* New York: Cambridge University Press.

Bernhardt, B.H., & Stemberger, J.P. (1998), *Handbook of phonological development: From the perspective of constraint-based nonlinear phonology.* San Diego: Academic Press.

Bernstein-Ratner, N. (1987). The phonology of parent child speech. In K.Nelson & A.van Kleeck (Eds.), *Children's Language: Vol. 6.* Hillsdale, NJ: Lawrence Erlbaum Associates.

——(1994). Phonological analysis of child speech. In J.L.Sokolov & C.E.Snow (Eds.), *Handbook of research in language development using CHILDES* (pp. 324– 372). Hillsdale, NJ: Lawrence Erlbaum Associates.

Blevins, J. (1995). The syllable in phonological theory. In J.Goldsmith (Ed.), *The handbook of phonological theory* (pp. 206–244). Oxford: Basil Blackwell.

Blumstein, S. (1973). *A phonological investigation of aphasic speech.* The Hague: Mouton.

Bod, R. (2001, January). *Introduction to probability theory in linguistics.* Paper presented at the Annual Meeting of the Linguistics Society of America, Washington, DC.

Boersma, P. (1997). How we learn variation, optionality, and probability. *Institute of Phonetic Sciences, University of Amsterdam, Proceedings 21,* 43–58.

——(1998). *Functional Phonology. Formalizing the interactions between articulatory and perceptual drives.* Doctoral dissertation, University of Amsterdam. The Hague: Holland Academic Graphics.

——(1999). Optimality-Theoretic Learning in the Praat Program. *Institute of Phonetic Sciences, University of Amsterdam, Proceedings 23,* 17–35.

Boersma, P., & Levelt, C. (2000). Gradual Constraint-ranking Learning Algorithm predicts acquisition order. In E.V.Clark (Ed.), *Proceedings of the Thirtieth Child Language Research Forum* (pp. 229–237). Stanford, CA: Center for the Study of Language and Information.

Booij, G., & Rubach, J. (1984). Morphological and prosodic domains in lexical phonology. *Phonology Yearbook 1,* 1–27.

Borowsky, T. (1986). *Topics in the lexical phonology of English.* Unpublished doctoral dissertation, University of Massachusetts at Amherst.

Botma, B., & Torre, E.J. van der (2000). The prosodic interpretation of sonorants in Dutch. In H.de Hoop & T.van der Wouden (Eds.), *Linguistics in the Netherlands 2000* (pp. 17–30). Amsterdam: John Benjamins.

Bradley, D. (1994). *A dictionary of the northern dialect of Lisu.* Canberra, Australia: Australian National University.

Braine, M.D.S. (1974). On what might constitute learnable phonology. *Language 50,* 270–299.
Bray, D. (1986). *The Brahui language.* Delhi, India: Gian Publishing House.
Brown, R. (1973). *A first language: The early stages.* Cambridge, MA: Harvard University Press.
Camarata, S., & Leonard, L. (1985). Children's production of nouns and verbs: Evidence for a semantic phonological interaction. *Papers and Reports in Child Language, 24,* 38–45.
——(1986). Young children pronounce object words more accurately than action words. *Journal of Child Language, 13,* 51–65.
Camarata, S., & Schwartz, R. (1985). Production of object words and action words: Evidence for a relationship between phonology and semantics. *Journal of Speech and Hearing Research, 28,* 323–330.
Carlson-Luden, V. (1979). *Causal understanding in the 10-month-old.* Unpublished doctoral dissertation, University of Colorado at Boulder.
Cassidy, K.W., & Kelly, M.H. (1991). Phonological information for grammatical category assignments. *Journal of Memory and Language, 20,* 348–369.
Chomsky, N. (1981). *Lectures on Government and Binding.* Dordrecht: Foris.
Clark, E. (1979). Building a vocabulary: Words for objects, actions and relations. In P.Fletcher & M.Garman (Eds.), *Language acquisition: Studies in first language development. New* York: Cambridge University Press.
Clements, G.N. (1990). The role of sonority cycle in core syllabification. In J.Kingston and M.Beckman (Eds.), *Papers in laboratory phonology: Vol I, Between the grammar and physics of speech* (pp. 283–333). Cambridge: Cambridge University Press.
Clements, G.N., & Keyser, S.J. (1983). *CV phonology: A generative theory of the syllable.* Cambridge, MA: MIT Press.
Coberly, M.S. (1985). *Initial and final consonant preferences across languages, in babbling, and in children's early speech.* Unpublished doctoral dissertation, University of Colorado.
Coleman, J., & Pierrehumbert, J. (1997). *Stochastic phonological grammars and acceptability.* Unpublished manuscript, Oxford University, Northwestern University, and Département Signal, ENST, Paris.
Craig, C.G. (1977). *The structure of Jacaltec.* Austin, TX: University of Texas Press.
Curtin, S. (2001). *Representational richness in phonological development.* Unpublished doctoral dissertation, University of Southern California.
——(in press). Children's Early Representations: Evidence from production and perception. To appear in the *Proceedings of the Holland Institute of Linguistics (HILP 5),* Potsdam, Germany.
Curtin, S., & Zuraw, K. (2001). *Learning the Ranking: Accounting for Variation in Acquisition.* Paper presented at Southwestern Workshop on Optimality Theory (SWOT) 6, University of Southern California, CA.
Danylenko, A., & Vakulenko, S. (1994). *Ukrainian.* Munich, Germany: Lincom Europa.
Dell, F. (1981). On the Learnability of Optional Phonological Rules. *Linguistic Inquiry, 12*(l), 31–37.
Demetras, M. (1989a, November). *Changes in parents' conversational responses: A function of grammatical development.* Paper presented at the American SpeechLanguage-Hearing Association (ASHA) Annual Conference, St. Louis, MO.
——(1989b). *Working parents conversational responses to their two-year-old sons.* Unpublished manuscript, University of Arizona, Tucson, AZ.
Demuth, K. (1995) Markedness and the development of prosodic structure. In J.Beckman (Ed.), *Proceedings of the North East Linguistic Society, 25,* 13–25. Amherst, MA: GLSA, University of Massachusetts.
——(2001). Prosodic constraints on morphological development. In J.Weissenborn & B.Höhle (Eds.), *Approaches to bootstrapping: Phonological, syntactic, and neurophysiological aspects of early language acquisition: Vol.2* (pp. 2–21). Amsterdam: John Benjamins.
Demuth, K., & Fee, E.J. (1995). *Minimal prosodic words in early phonological development.* Unpublished manuscript, Brown University and Dalhousie University.
Derwing, B.L. (1992). A 'Pause-Break' task for eliciting syllable boundary judgements from literate and illiterate speakers: Preliminary results from five diverse languages. *Language and Speech, 35,* 219–235.
Dewey, G. (1923). *Relative frequency of English speech sounds.* Cambridge, MA: Harvard University Press.

Dinnsen, D.A. (1996). Context effects in the acquisition of fricatives. In B.Bernhardt, J.Gilbert & D.Ingram (Eds.), *Proceedings of the UBC International Conference on Phonological Acquisition* (pp. 136–148). Somerville, MA: Cascadilla Press.

Dobrich, W., & Scarborough, H.S. (1992). Phonological characteristics of words young children try to say. *Journal of Child Language, 19,* 597–616.

Doniach, N.S., & Kahane, A. (1996). *The Oxford English-Hebrew dictionary.* New York: Oxford University Press.

Eckman, F.R., & Iverson, G.K. (1994). Pronunciation difficulties in ESL: Coda consonants in English interlanguage. In M.Yavas (Ed.), *First and second language phonology* (pp. 251–266). San Diego: Singular Publishing Group.

Elbert, S, & Pukui, M.K. (1979). *Hawaiian grammar.* Honolulu, HI: University Press of Hawaii.

England, N.C. (1983). *A grammar of Mam, a Mayan language.* Austin, TX: University of Texas Press.

Fernández, Z. (1996). *Pima Bajo.* Munich, Germany: Lincom Europa.

Faingold, E.D. (1990). The acquisition of syllabic and word structure: Individual differences and universal constraints. *Language Sciences, 12*(1), 101–113.

Fallows, D. (1981). Experimental evidence for English syllabification and syllable structure. *Journal of Linguistics, 17,* 309–317.

Fee, E.J., & Ingram, D. (1982). Reduplication as a strategy of phonological development. *Journal of Child Language, 2,* 198–220.

Fennell, T.G., & Gelsen, H. (1980). *A grammar of modern Latvian: Vol. 3.* The Hague: Mouton.

Fenson, L., Dale, P., Reznick, S., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J. (1993). *MacArthur Communicative Development Inventories.* San Diego, CA: Singular Publishing Group, Inc.

Ferguson, C.A. (1963). Assumptions about nasals; a sample study in phonological universals. In J.Greenberg (Ed.), *Universals of language* (pp. 53–60). Cambridge, MA: MIT Press.

——(1978). Fricatives in child language acquisition. In V.Honsa & M.H.HardmanBautista (Eds.), *Papers on Linguistics and Child Language* (pp. 61–85). The Hague: Mouton.

Fikkert, P. (1994). *On the acquisition of prosodic structure.* Leiden: Holland Institute of Generative Linguistics.

Flexner, S.B. (Ed.). (1987). *The Random House dictionary of the English language* (2nd ed.). New York: Random House.

Foley, J. (1972). Rule precursors and phonological change by meta-rule. In R.Stockwell & R.Macaulay (Eds.), *Linguistic change and generative theory.* Bloomington, IN: Indiana University Press.

Fonte, I. (1996). *Restrictions on codas: An Optimality Theoretic account of phonotactics.* Unpublished masters thesis. McGill University, Montreal, Canada.

Franke, C. (1912). Über die erste Lautstufe der Kinder (C.Murchison & S.Langer, Trans.). In A.Bar-Adon & W.F.Leopold (Eds.), *Child language: A book of readings* (pp. 13–17). Englewood Cliffs, NJ: Prentice-Hall. (Reprinted from Tiedemann's Observations on the development of the mental faculties of children, by C.Murchison & S.K.Langer, 1927, *The Pedagogical Seminary and Journal of Genetic Psychology 34,* 205–230)

French, A. (1989). The systematic acquisition of word forms by a child during the firstfifty word stage. *Journal of Child Language, 16,* 69–90.

Garrett, J., Lastowka, G., Naahielua, K., & Pallipamu, M. (Eds.). (1996). *TurkmenEnglish dictionary* [On-line]. Internet: http://www.chaihana.com//dict.doc

Gathercole, S.E. (1995). Is nonword repetition a test of phonological memory or longterm knowledge? It all depends on the nonwords. *Memory & Cognition, 23*(1), 83– 94.

Gennari, S., & Demuth, K. (1997). Syllable Omission in the Acquisition of Spanish. In E.Hughes, M.Hughes & A.Greenhill (Ed.), *Proceedings of the Annual Boston University Conference on Language Development: Vol. 1* (pp. 182–193).

Gnanadesikan, A. (1995). *Markedness and faithfulness constraints in child phonology.* Unpublished manuscript, University of Massachusetts, Amherst. Rutgers Optimality Archive.

Goad, H. (1997a). Codas, word minimality, and empty-headed syllables. In E.V.Clark (Ed.), *Proceedings of the twenty-eighth annual child language research forum* (pp. 113–122). Stanford, CA: Center for the Study of Language and Information.

——(1997b). Consonant harmony in child language: An optimality-theoretic account. In S.J.Hannah & M.Young-Scholten (Eds.). *Generative studies in the acquisition of phonology* (pp. 113–142). Amsterdam: John Benjamins.

Goad, H., & Brannen, K. (in press). Syllabification at the right edge of words: Parallels between child and adult grammars. In J.van de Weijer (Ed.), *Issues in the phonology-phonetics interface.* Amsterdam: John Benjamins.

Goldstein, M.C., & Nornang, N. (1970). *Modern spoken Tibetan: Lhasa dialect.* Seattle, WA: University of Washington Press.

Gordon, M. (2000). *Syllable weight: phonetics, phonology, and typology.* Unpublished doctoral dissertation, UCLA.

Greenberg, J.H. (1954). A quantitative approach to the morphological typology of language. In R.F.Spencer (Ed.), *Method and perspective in anthropology* (pp. 192– 220). Minneapolis: University of Minnesota Press.

——(1963). Some universals of grammar with particular reference to the order of meaningful elements. in J.H.Greenberg (Ed.), *Universals of language* (pp. 58–90). Cambridge, MA: MIT Press.

——(1966). *Language universals, with specific reference to feature hierarchies.* The Hague: Mouton.

——(1974). *Language typology: A historical and analytic overview.* The Hague: Mouton.

——(1978). Some generalizations concerning initial and final consonant clusters. In J.H.Greenberg (Ed.), *Universals of human language: Vol. 2. Phonology* (pp. 243– 280). Stanford, CA: Stanford University Press.

Grimes, J.E., & Grimes, B.F. (1999). *Ethnologue: Languages of the world* [CD-ROM]. Dallas, TX: Summer Institute of Linguistics.

Gruzdeva, E. (1998). *Nivkh.* Munich, Germany: Lincom Europa.

Hale, M., & Reiss, C. (1995). *The initial ranking of faithfulness consonants in UG.* Unpublished manuscript: Harvard University and Concordia University. Rutgers Optimality Archive.

Halle, M., & Vergnaud, J-R. (1980). Three dimensional phonology. *Journal of Linguistics Research, 1,* 83–105.

Hammond, M. (1997). Vowel quantity and syllabification in English. *Language, 73,* 1– 17.

——(1999). *The phonology of English.* New York: Oxford University Press.

——(2001, April). *English phonotactics and probabilistic phonology.* Paper presented at Southwestern Workshop on Optimality Theory (SWOT) 6, University of Southern California, CA.

Hammond, M., Moravcsik, E.A., & Wirth, J.R. (1988). Language typology and linguistic explanation. In M.Hammond, E.A.Moravcsik, & J.R.Wirth (Eds.), *Studies in Syntactic Typology Vol. 17* (pp. 1–22). Amsterdam: John Benjamins.

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

——(1997). Licensing inheritance: An integrated theory of neutralisation. *Phonology, 7,* 315–370.

Harris, M., Barrett, M., Jones, D., & Brookes, S. (1988). Linguistic input and early word meaning. *Journal of Child Language, 15*(1), 77–94.

Hart, B. (1991). Input frequency and children's first words. *First Language, 11,* 289–300.

Hayes, B., & MacEachern, M. (1998). Quatrain form in English folk verse. *Language, 73*(3) ,473–507.

Headley, R.K. Jr., Chhor, K., Lim, L.K., Cheang, L.H., & Chun, C. (1977). *CambodianEnglish dictionary* (Vol. 1–2). Washington: Catholic University of American Press.

Healey, A. (1977). *Telefol phonology.* Canberra, Australia: Australian National University.

Healey, P.M., & Healey, A. (1977). *Telefol dictionary.* Canberra, Australia: Australian National University.

Higginson, R.P. (1985). *Fixing-assimilation in language acquisition.* Unpublished doctoral dissertation. Washington State University.

Holmes, U.T. (1927). The phonology of an English-speaking child. *American Speech, 2* (5), 212–225.

Hualde, J.I. (1992). *Catalan.* New York: Routledge.

Huang, P.P. (1970). *Cantonese dictionary: Cantonese-English, English-Cantonese.* New Haven, CT: Yale University Press.

Hyams, N.M. (1986). *Language acquisition and the theory of parameters.* Dordrecht: Reidel.

Ingram, D. (1974). Phonological rules in young children. *Journal of Child language, 1,* 49–64.

——(1988) The acquisition of word-initial /v/. *Language and Speech, 31*(1), 77–85.

Irwin, B. (1974). *Salt-Yui grammar.* Canberra, Australia: Australian National University.

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

Jakobson, R. (1968). *Child Language, aphasia, and phonological universals.* The Hague: Mouton. (Original work published in 1941).

Jespersen, O. (1904). *Lehrbuch der Phonetik* (H. Davidson, Trans.). Leipzig and Berlin, Germany: B.G.Teubner.

Jusczyk, P.Q., Luce, P.A., & Charles-Luce, J. (1994). Infant's sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language, 33,* 630–645.

Kahn, D. (1976). *Syllable-based generalizations in English phonology.* Bloomington, IN: Indiana University, Linguistics Club.

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

Keegan, J.M. (1997). *A reference grammar of Mbay.* Munich, Germany: Lincom Europa.

Keegan, J.M., Marcel, N., & Bertin, M.T. (1996). *Dictionary of Mbay.* Munich, Germany: Lincom Europa.

Kehoe, M.M., & Stoel-Gammon, C. (2001). Development of syllable structure in English-speaking children with particular reference to rhymes. *Journal of Child Language, 28,* 393–433

Kelly, M.H. (1996). The role of phonology in grammatical category assignments. In J.L. Morgan & K.Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 249–262). Mahwah, N.J.: Lawrence Erlbaum Associates.

Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language, 37,* 295–311.

Key, H. (1967). *Morphology of Cayuvava.* The Hague: Mouton.

Kozintseva, N.A. (1998). *Modern Eastern Armenian.* Munich, Germany: Lincom Europa.

Krámský, J. (1946–1948). Fonogické využiti samohláskovych fonemát. *Linguisticia Slovaca* IV, VI.

——(1959a). A quantitative typology of languages, *L & S,* 72–85.

——(1959b). Fonologické využiti samohláskovych fonemát, *Linguistica Slovaca* IV, VI.

Ladefoged, P. (1993). *A course in phonetics.* New York: Harcourt Brace College Publishers.

Leopold, W.F. (1939). *Speech development of a bilingual child, a linguist's record: Vol. 1. Vocabulary Growth in the First Two Years.* New York: AMS Press.

Leuninger, H. (1983). Markedness, modularity, and functional explanations. In S.W.Felix, H.Wode (Eds.), *Language development and the crossroads: papers from the Interdisciplinary Conference on Language Acquisition at Passau* (pp. 31– 48). Tübingen, Germany: Narr.

Levelt, C., Schiller, N., & Levelt, W. (2000). The acquisition of syllable types. *Language Acquisition, 8*(3), 237–264.

Levelt, C., & Vijver, R.van de (in press). The acquisition of syllable types in crosslinguistic and developmental grammars. In R.Kager, J.Pater & W.Zonneveld (Eds.), *Fixing Priorities: Constraints in Phonological Acquisition.* Cambridge: Cambridge University Press.

Lewis, G.L. (1967). *Turkish grammar.* Cambridge: Cambridge University Press.

Lichtenberk, F. (1983). *A grammar of Manam.* Honolulu, HI: University of Hawaii Press.

Lleó, C., & Demuth, K. (1998). Prosodic constraints on the emergence of grammatical morphemes: Crosslinguistic evidence from Germanic and Romance languages. In A.Greenhill, H.Litterfield & C.Tano (Eds.), *Proceedings of the 23rd Boston University Conference on Language Development: Vol 2* (pp.407–418). Somerville, MA: Cascadilla Press.

Lombardi, L. (1991). *Laryngeal features and laryngeal neutralization.* New York: Garland Press.

——(1996). *Why place and voice are different: Constraint interactions and featural faithfulness in Optimality Theory.* Rutgers Optimality Archive.

Luo, Y. (1998). *A dictionary of Dehong, Southwest China.* Canberra, Australia: Pacific Linguistics.

MacKay, C.J. (1999). *A grammar of Misantla Totonac.* Salt Lake City, UT: University of Utah Press.

Macken, M. (1995). Phonological acquisition. In J.Goldsmith (Ed.), *The handbook of phonological theory* (pp. 671–696). Oxford: Basil Blackwell.

Macken, M., & Barton, D. (1980). A longitudinal study of the acquisition of the voicing contrast in American-English word-initial stops, as measured by voice onset time. *Journal of Child Language, 7,* 41–74.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
——(Ed.). (1999). *The emergence of language.* Mahwah, NJ: Lawrence Erlbaum Associates.
Maddieson, I. (1984). *Patterns of Sounds.* Cambridge, MA: Cambridge University Press.
Manich Jumsai, M.L. (1958). *Thai-English dictionary* (48th ed.). Bangkok, Thailand: Chalermnit.
Marchman, V., & Plunkett, K. (1989). Token frequency and phonological predictability in a pattern association network: Implications for child language acquisition. *Proceedings of the Eleventh Annual Meeting of the Cognitive Science Society* (pp. 179–187). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
McCarthy, J.J. (1979). *Formal problems in Semitic morphology and phonology.* Doctoral dissertation, MIT. New York: Garland Press, 1985.
McCarthy, J.J., & Prince, A. (1993). *Prosodic Morphology I: Constraint Interaction and Satisfaction.* Unpublished manuscript, University of Massachusetts and Rutgers University.
McCarus, E.N. (1967). *Kurdish-English dictionary, dialect of Sulaimania, Iraq.* Ann Arbor, MI: University of Michigan Press.
Menn, L. (1971). Phonotactic rules in beginning speech. *Lingua, 26,* 225–251.
Messer, S. (1967). Implicit phonology in children. *Journal of Verbal Learning and Verbal Behavior, 6,* 609–613.
Mines, M.A., Hanson, B.F., & Shoup, J.E. (1978). Frequency of occurrence of phonemes in conversational English. *Language and Speech, 21*(3), 221–241.
Morén, B. (1998). The puzzle of Kashmiri stress. In H.Fukazawa, F., *University of Maryland Working Papers in Linguistics, 7,* 126–149.
Morgan, J.L. (1990). Input, innateness, and induction in language acquisition. *Developmental Psychobiology, 23*(3), 661:678.
Moskowitz, A.I. (1970). The two-year-old stage in the acquisition of English phonology. *Language, 46,* 426–441.
Mowrer, D., & Burger, S. (1991). A comparative analysis of phonological acquisition of consonants in the speech of 2 ½-6-year-old Xhosa- and English-speaking children. *Clinical Linguistics and Phonetics, 5*(2), 139–164.
Murane, E. (1974). *Daga grammar, from morpheme to discourse.* Norman, OK: Summer Institute of Linguistics of the University of Oklahoma.
Naigles, L., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language, 25,* 95–120.
Nau, N. (1998). *Latvian.* Munich, Germany: Lincom Europa.
Niang, M.O. (1997). *Constraints on Pulaar phonology.* London: University Press of America.
Ohala, D. (1992). A unified theory of final consonant deletion in early child speech. In C.Gerfen & P.Pinar (Eds.), *Coyote Papers: Working Papers in Linguistics from A → Z* (pp. 96–117). Tucson, AZ: University of Arizona.
——(1994, February). *Constraints and syllables in acquisition.* Talk presented at the Experimental Linguistics Group, University of Arizona, Tucson, AZ.
——(1996). *Cluster reduction and constraints in acquisition.* Unpublished doctoral dissertation, University of Arizona, Tucson, AZ.
Olmsted, D.L. (1966). A theory of the child's learning of phonology. *Language, 42 (2),* 531–535.
——(1971). *Out of the mouths of babes: Earliest stages in language learning.* The Hague: Mouton.
Ouden, D.B. den & Bastiaanse, R. (2001). *Segmental vs. syllable markedness: Deletion errors in the paraphasias of fluent and nonfluent aphasics.* Unpublished manuscript, University of Groningen, Groningen, The Netherlands.
Paradis, C., & Prunet, J. (Eds.). (1991). *The special status of coronals: Internal and external evidence: Phonology and Phonetics 2.* San Diego, CA: Academic Press.
Perkins, R.D. (1989). Statistical techniques for determining language sample size. *Studies in Language, 13*(2), 293–315.
Pierrehumbert, J. (1994). Syllable structure and word structure: A study of triconsonantal clusters in English. In P.A.Keating (Ed.), *Phonological structure and phonetic form: Papers in Laboratory Phonology III* (pp. 168–188). Cambridge: Cambridge University Press.
Piggott, G.L. (1991). Apocope and the licensing of empty-headed syllables. *The Linguistic Review, 8,* 287–318.
——(1999). At the right edge of words. *The Linguistic Review, 16,* 143–185.

——(in press). The phonotactics of a 'Prince' language: A case study. *The Linguistic Review.*

Poole, I. (1934). Genetic development of articulation of consonant sounds in speech. *Elementary English Review 11,* 159–161.

Prince, A., & Smolensky, P. (1993). *Optimality Theory: Constraint interaction in generative grammar.* RuCCs Technical Report #2, Rutgers University Center for Cognitive Science.

Pye, C., Ingram, D., & List, H. (1987). A comparison of initial consonant acquisition in English and Quiché. In K.E.Nelson & A.Van Kleech (Eds.), *Children's language:Vol. 6* (pp.175–190). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rehg, K., & Sohl, D. (1981). *Ponapean Reference Grammar.* Honolulu, HI: University Press of Hawaii.

Rice, K. (1993). Default variability: The coronal-velar relationship. *Toronto Working Papers in Linguistics, 12*(2), 97–130.

Roark, B., & Demuth K. (2000). Prosodic constraints and the learner's environment: A corpus study. In S.C.Howell, S.A.Fish & T.Keith-Lucas (Eds.), *Proceedings of the 24th Boston University Conference on Language Development: Vol. 2* (pp. 597– 608). Somerville, MA: Cascadilla Press.

Roberts, A.H. (1965). *A statistical linguistic analysis of American English.* The Hague: Mouton.

Salidis, J., & Johnson, J.S. (1997). The production of minimal words: A longitudinal case study of phonological development. *Language Acquisition, 6*(10), 1–36.

Scarry, R. (1973). *Best word book ever* (15th ed.). New York: Western Publishing.

——(1994). *Busy, busy town.* New York: Western Publishing.

Scobbie, J.M., Gibbon, F., Hardcastle, W.J., & Fletcher, P. (2000). Covert contrast as a stage in the acquisition of phonetics and phonology. In M.Broe & J.Pierrehumbert (Eds.), *Papers in Laboratory Phonology V: Language Acquisition and the Lexicon* (pp. 194–207). Cambridge: Cambridge University Press.

Seegmiller, S. (1996). *Karachay.* Munich, Germany: Lincom Europa.

Selkirk, E. (1982). Syllables. In H.van der Hulst & N.Smith (Ed.), *The structure of phonological representations* (pp. 337–383). Dordrecht: Foris.

——(1984). On the major class features and syllable theory. In M.Arnoff & R.T.Oehrle (Eds.), *Language sound structure* (pp. 107–136). Cambridge, MA: MIT Press.

Sereno, J.A., & Jongman, A. (1990). Phonological and form class relations in the lexicon. *Journal of Psycholinguistic Research, 19*(6), 387–404

Sharpe, M.C. (1972). *Alawa phonology and grammar.* Canberra, Australia: Australian Institute of Aboriginal Studies.

Sherman, D. (1975). Noun-verb stress alternation: An example of the lexical diffusion of sound change in English. *Linguistics, 149,* 43–71.

Shibamoto, J.S., & Olmsted, D.L. (1978). Lexical and syllabic patterns in phonological acquisition. *Journal of Child Language, 5,* 417–446.

Sievers, E. (1881). *Grundzüge der Phonetik.* Leipzig, Germany: Breitkopf and Hartel.

Sinha, N.K. (1975). *Mundari grammar.* Mysore, India: Central Institute of Indian Languages.

Smit, A.B., Hand, L., Freilinger, J.J., Bernthal, J.E., & Bird, A. (1990) The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55,* 779–798.

Smith, B.L. (1979). A phonetic analysis of consonantal devoicing in children's speech. *Journal of Child Language, 6,* 19–28.

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

Smolensky, P. (1996). On the comprehension/production dilemma in child language. *Linguistic Inquiry, 27,* 720–731.

Snow, K. (1963). A detailed analysis of articulation responses of "normal" first grade children. *Journal of Speech and Hearing Research, 6,* 277–290.

Spencer, A. (1996). *Phonology.* Cambridge, MA: Blackwell Publishers.

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

Steriade, D. (1982). *Greek prosodies and the nature of syllabification.* Unpublished doctoral dissertation, MIT, Cambridge, MA.

Stoel-Gammon, C. (1985). Phonetic inventories, 15–24 months: A longitudinal study. *Journal of Speech & Hearing Research, 28,* 505–512.

——(1995). *What children say: A phonological analysis of the target words acquired by young American children.* Paper presented at Child Language Seminar, Bristol University, England.

——(1998). Sounds and words in early language acquisition: The relationship between lexical and phonological development. In R.Paul (Ed.) *Exploring the Speech-Language Connection* (pp. 25–52). Baltimore: Paul H.Brookes Publishing Co.

Storkel, H.L. (1999). Sublexical influences on lexical development in children. *Research on Spoken Language Processing. Progress Report No. 23* (pp. 211–225) Bloomington, IN: Speech Research Laboratory Indiana University.

——(2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research, 44,* 1231–1337.

Tesar, B. (1995). *Computation Optimality Theory.* Unpublished doctoral dissertation, University of Colorado, Boulder. Rutgers Optimality Archive.

Tesar, B., & Smolensky, P. (1993). The learnability of Optimality Theory: An algorithm and some basic complexity results. *Technical Report CU-CS-67S-93.* Department of Computer Sciences, University of Colorado, Boulder. Rutgers Optimality Archive.

——(1998). Learnability in Optimality Theory. *Linguistic Inquiry, 29,* 229–268.

Tharp, J.A., & Y-Bham D. (1980). *A Rhade-English dictionary with English-Rhade finderlist.* Canberra, Australia: Australian National University.

Thurber, J. (1961). The Last Clock. In *Lanterns and Lances.* New York: Harper & Row.

Treiman, R., Kessler, B., Knewasser, S., Tincoff, R., & Bowman, M. (2000). English speakers' sensitivity to phonotactic patterns. In J.Pierrehumbert & M.Broe (Eds.), *Papers in Laboratory Phonology V: Language acquisition and the lexicon* (pp. 269– 282). Cambridge: Cambridge University Press.

Treiman, R., & Danis, C. (1988). Syllabification of intervocalic consonants. *Journal of Memory and Language, 27,* 87–104.

Trubetzkoy, N.S. (1931). Die phonologischen Systeme. *Travaux du Cercle Linguistique de Prague, 4,* 96–116.

——(1969). *Principles of Phonology* (C.A.M. Baltaxe, Trans.). Berkeley, CA: University of California Press. (Original work published 1939).

Tryon, D.T. (1967a). *Dehu Grammar.* Canberra, Australia: Australian National University.

——(1967b). *Dehu-English Dictionary.* Canberra, Australia: Australian National University.

Turkina, E. (1982). *Latviesu-anglu vardnica: Ap 30,000 vardu.* Riga, Latvia: Avots.

Valdman, A., Klinger, T.A., Marshell, M.M., & Roltet, K.J. (1988). *Dictionary of Louisiana Creole.* Bloomington, IN: Indiana University Press.

Velleman, S.L. (2000, January). *The Optimal Initial State.* Paper presented at the Annual Meeting of the Linguistics Society of America, Chicago IL.

Velten, H. (1943) The growth of phonemic and lexical patterns in infant language. *Language, 19,* 281–292.

Vihman, M.M. (1992). Early syllables and the construction of phonology. In C.A. Ferguson, L.Menn & C.Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (393–422). Timonium, MD: York Press.

——(1996). *Phonological development: The origins of language in the child.* Cambridge, MA: Blackwell.

Vihman, M.M., & Boysson-Bardies, B.de (1994). The nature and origins of ambient language influences on infant vocal production and early words. *Phonetica, 51,* 159– 69

Vihman, M.M., & Ferguson, C.A. (1987). The acquisition of final consonants. In *Proceedings of the Eleventh International Congress of Phonetic Sciences* (pp. 381– 384). Tallinn, Estonia, USSR.

Vihman, M.M., & McCune, L. (1994). When is a word a word? *Journal of Child Language 21,* 517–542.

Vihman, M.M., & Velleman, S.L. (1989). Phonological reorganization: A case study. *Language and Speech, 32,* 149–170.

Vihman, M.M., Velleman, S.L., & McCune, L. (1994). How abstract is child phonology? Towards an integration of linguistic and psychological approaches. In M.Yavas, (Ed.), *First and second language phonology* (pp. 9–44). San Diego: Singular Publishing Group.

Walsh, D.S., & Biggs, B. (1966). *Proto-Polynesian word list I.* Auckland, New Zealand: Linguistics Society of New Zealand.

Warren-Leubecker, A. (1982). *Sex differences in speech to children.* Georgia Institute of Technology.

Warren-Leubecker, A., & Bohannon, J.N. (1984). Intonation patterns in child-directed speech: Mother-father speech. *Child Development, 55,* 1379–1385.

Webster, M. (1964). *The new Merriam-Webster pocket dictionary.* New York: Simon and Schusterp.

Welleman, B.L., Case, I.M., Mengert, K.G., & Bradbury, D.E. (1931). Speech sounds of young children. *University of Iowa Studies in Child Welfare: Vol. V, no. 2.* Iowa: Iowa University City.

Werker, J.F., & Tees, R.C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7,* 49–63.

Wiering, E., & Wiering, M. (1994). *The Doyayo language.* Dallas, TX: University of Texas at Arlington.

Wilkinson, K. (1988). Prosodic structure and Lardil phonology. *Linguistic Inquiry, 19,* 325–334.

Wu, C. (1998). *Daur.* Munich, Germany: Lincom Europa.

Yip, M. (1991). Coronals, consonants clusters and the coda condition. In C.Paradis & J. Prunet (Eds.), *The special status of coronals: Internal and external evidence: Phonology and Phonetics 2* (pp. 61–78). San Diego, CA: Academic Press.

Zamuner, T.S. (1996). *The Acquisition of coda consonants: Case study of Smith (1973).* Unpublished manuscript, McGill University, Montreal, Canada.

——(1997). *Distributions of word final codas in Iai.* Unpublished manuscript, University of Arizona, Tucson, AZ.

——(1998a). *Regularity effects on the distribution of word-final consonants in English trochaic and iambic words.* Unpublished manuscript, University of Arizona, Tucson, AZ.

——(1998b). *Rethinking the role of Universal Grammar in language acquisition: Evidence from the distribution of English word-final coda consonants.* Unpublished manuscript, University of Arizona, Tucson, AZ.

Zamuner, T.S., & Gerken, L.A. (1997). Young children's production of coda consonants in different prosodic environments. In E.V.Clark (Ed.), *Proceedings of the twentyninth annual child language research forum* (pp. 13–25). Stanford, CA: Center for the Study of Language and Information.

Zamuner, T.S., & Ohala, D. (1999). Preliterate children's syllabification of intervocalic consonants. In A.Greenhill, H.Litterfield & C.Tano (Eds.), *Proceedings of the 23rd Boston University Conference on Language Development: Vol.2* (pp. 753– 763). Somerville, MA: Cascadilla Press.

Zec, D. (1988). *Sonority constraints on prosodic structure.* Doctoral dissertation, Stanford University. New York: Garland Press, 1994.

Zipf, G.K. (1935). *Psychobiology of language: An introduction to dynamic philology.* Cambridge, MA: M.I.T. Press (1965).

Zoll, C. (1998). *Positional asymmetries and licensing.* Unpublished doctoral dissertation, MIT. Rutgers Optimality Archive.

Language Index

Author Index

Subject Index