Acquisition of Segmental Structure: Consequences for Speech Perception and Second Language Acquisition

Cynthia A. Brown

Department of Linguistics
McGill University
Montréal, Québec

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

September 1997

0-612-44371-X

"Are you an Ephraimite?"
When he said, "No," they said to him
"Then say Shibboleth,"
and he said,
"Sibboleth"
for he could not pronounce it right;
then they seized him and slew him at the fords of the Jordan.
And there fell at that time forty-two thousand of the Ephraimites.

Judges 12:5-6

Thankfully, the consequences of L1 phonological interference are not so grave today....

Table of Contents

Abstract

Through an investigation of the acquisition of feature geometric representations in first and second language acquisition, this dissertation demonstrates how the Feature Geometry theory contained in Universal Grammar actively guides and constrains the acquisition of segmental representations by children. In addition, it demonstrates how the mature feature geometry in a speaker's mental grammar restricts the range of non-native phonemic contrasts that he or she will be sensitive to in the input and, hence, able to acquire as an L2 learner.

Three related areas of research are explored and integrated in this work: first, a theoretical study explores the feature-geometric representation of sonorant and non-sonorant laterals, based on their behavior in a variety of phonological processes cross-linguistically, and suggests that [lateral] is not a phonological feature, but rather that laterality is a phonetic property that derives from a specific feature-geometric representation; second, an experimental study investigates the acquisition of phonemic contrasts by English children and demonstrates that segmental representations are acquired in a uniform order that is consistent with properties of Feature Geometry; finally, a series of experimental studies examines the perception and acquisition of the English /l-r/, /b-v/, /p-f/, /f-v/ and /s-θ/ contrasts by native speakers of Japanese, Mandarin Chinese and Korean.

The findings from each of these studies are synthesized to obtain a comprehensive picture of how segmental representations are acquired and how this L1 knowledge impinges on the acquisition of L2 phonemes: it is argued that the monotonic acquisition of feature-geometric structure by young children restricts their sensitivity to particular non-native contrasts, and the continued operation of this existing feature geometry in adult speech perception constrains which non-native contrasts adult learners will be sensitive to in the L2 input and, therefore, capable of acquiring; the circumstances in which the native grammar facilitates perception of non-native contrasts and in which acquisition is possible are also discussed.

Résumé

À travers l'étude de l'acquisition des représentations géométriques de traits en apprentissage d'une première et seconde langue, cette thèse démontre que la théorie de la Géométrie de Traits, qui relève de la Grammaire Universelle, guide et régit activement l'acquisition des représentations segmentales chez les enfants. De plus, il est montré que la géométrie de traits présente dans la grammaire maternelle d'un locuteur restreint le champ de contrastes phonémiques non natifs auxquels cette persone sera sensible et qu'elle sera donc à même d'acquérir en tant qu'apprenant d'une langue seconde.

Trois domaines de recherche sont examinés et reliés au sein de ce travail: tout d'abord, une étude théorique explore la représentation, en termes de géométrie de traits, des latérales sonorantes et non-sonorantes en se basant sur leur comportement dans plusieurs processus phonologiques à travers les langues. Il est suggèré que [latéral] n'est pas un trait phonologique mais que le caractère latéral est une propriété phonétique qui dérive d'une représentation géométrique de traits particulière. Ensuite, une étude expérimentale examine l'acquisition de contrastes phonémiques chez les enfants anglais et montre que les représentations segmentales sont acquises suivant un ordre uniforme qui est compatible avec la Géométrie de Traits. Enfin, une série d'études expérimentales explore la perception et l'acquisition des contrastes anglais /l-r/, /b-v/, /p-f/, /f-v/ et /s-θ/ par des apprenants dont la langue maternelle est le japonais, le chinois mandarin et le coréen.

Les résultats de chacune de ces études sont synthétisés de façon à obtenir une image complète de la façon dont les représentations segmentales sont acquises et dont la connaissance native affecte l'acquisition des phonèmes de la L2. Nous montrons que l'acquisition monotonique de la structure géométrique de traits par les jeunes enfants restreint leur sensibilité à certains contrates non natifs et que dans la perception adulte l'opération continuelle de la géométrie de traits déjà existente régit la nature des contrates non natifs auxquels les apprenants adultes seront sensibles dans la langue

seconde et qu'ils seront donc capables d'acquérir. Les circonstances dans lesquelles la grammaire maternelle facilite la perception de contrastes non natifs et où de nouvelles représentations segmentales sont acquises avec succès sont également discutées.

Acknowledgements for the Thesis

Anyone who has completed a dissertation knows that it is a long climb up a sometimes steep mountain. How can one, in just a few short pages, thank all of those who have served either as inspiration, guide, fellow traveler, or member of the base camp during that long journey. No one makes the trip alone. I would like to take this opportunity (as I approach the summit) to acknowledge and thank those individuals that have helped me along the way, through their example, guidance and support.

My first debt of gratitude is owed to my supervisors, Lydia White and Heather Goad. I have been fortunate to have had two supervisors for the writing of this dissertation (perhaps doubly blessed because they rarely, if ever, disagreed in their direction). Given the large number of students Lydia has supervised in recent years, it is very difficult (if not impossible) to say anything original. Each and every acknowledgements attests to her keen ability to draw out the implications of any new twist of linguistic theory for language acquisition, her direct no-nonsense approach which (thankfully) manages to pare down any and all superfluous prose, and her seemingly supernatural efficiency which is now legendary among McGill students. This acknowledgements is no exception; I thank her for lending all of these talents to the supervision of this work. Any clarity of organization and argumentation found in this thesis are due to her. (If only I had learned her trick for squeezing 36 hours out of each day!)

Heather Goad joined the McGill linguistics department just as I was beginning to explore the implications of phonological theory for acquisition. It was without question her superb knowledge of both phonological theory and acquisition theory (and her enthusiasm) that enabled me to carry out the research in this dissertation. She has given a level of detailed attention to my work that most professors reserve for their own. I thank Heather for her thoroughness in commenting on all of my work, her ability to always find the time for "just one more" meeting, and for the professional advice she has given me throughout the time I've known her (I'm sure more than one abstract acceptance was due to her careful editing).

Though not my thesis advisor, Glyne Piggott has contributed tremendously to my linguistic education and to my conception of what a true scholar is. More than anyone, he has given me the courage to follow my linguistic instincts, even when that means going against the current theoretical tide. Many of my linguistic "highs" came during the hour after his seminars, when I would scribble away on the chalkboard trying to prove that his analysis just couldn't be right; he encouraged that kind of challenge to his ideas (an attitude which is vital to scientific progress, but so often sadly absent). I must also thank Glyne for cultivating my taste for Indian food, though I'm not so sure I agree with him that "if you don't sweat, you're not enjoying it!"

I am also grateful to the other members of the Linguistics department, Prof. Mark Baker, Prof. Nicole Domingue, Prof. Nigel Duffield, Prof. Brendan Gillon, Prof. Myrna Gopnik, Prof. Michel Paradis and Prof. Lisa Travis, for their contributions to my education and development. Each had something to teach me, if not always directly related to my research – how to raise spider plants, that one must make time and space for love, that summers are for working (!), and that our current theories of grammar are probably not correct (but that that's O.K., it only means there's a lot more research to do). Thank you for creating a stimulating yet supportive atmosphere in which I could learn and develop my linguistic skills.

My appreciation must also be expressed to Lise Vinette and Linda Suen, the department administrators, who do far more for students than we realize, or thank them for. More than once, I have depended on their patient help and vast knowledge of university procedure to navigate my way through some intricacy of the graduate program.

Much of what I have learned about doing experimental research I learned through my involvement in Lydia White's research projects. Her respect for students' ideas and her inclusion of her research assistants in every step of a research project has been a tremendous aid in the development of my own research program. It has been my privilege to have worked on several different projects with Susan Bennett, Joyce Bruhn de Garavito, Dongdong Chen, Makiko Hirakawa, Alan Juffs, Takako Kawasaki, Silvina Montrul, Joe Pater and Philippe Prévost. I am grateful for having had the opportunity to learn about L2 syntactic acquisition from them and I appreciate their willing support of my phonological research.

Anyone who does experimental work knows that sometimes the most challenging aspect of a project is locating adequate numbers of appropriate subjects. This task was facilitated tremendously by the generous assistance of Takako Kawasaki, Makiko Hirakawa and Dongdong Chen who recruited Japanese and Chinese subjects for me in Montreal, Canada, and by Michie Namita who found Korean and Chinese subject for me in Sapporo, Japan.

Among my many fellow "journeyers", both past and present, I would like to thank Shanley Allen, Mengistu Amberber, Doris Belland, Kathy Burns, Ted Caldwell, Jenny Dalalakis, Suzy Fukuda, Shinji Fukuda, Marie Hannan, Yahiro Hirakawa, Martine Kessler Robb, Alan Libert, John Matthews, Amanda Miller, Vicky Murphy, Masanori Nakamura, Jennifer Ormston, Johanne Paradis, Dominique Rodier, Sharon Rose, Ben Shaer and Roumyana Slabakova. I would especially like to thank Marie Hannan and Ted Caldwell for sharing their lives with me and for forcing me away from my desk more than once to get out and have some fun.

My work and development as a linguist has also been aided by conversations I have had along the way with the following people: John Archibald, Catherine Best, Abby Cohn, Stephen Crain, Jane Fee, Paula Fikkert, Phil Hamilton, Barbara Hancin-Bhatt, Henrietta Hung, Keith Johnson, Claartje Levelt, Jaye Padgett, Barbara Pearson, Keren Rice, Jeroen van de Weijer and Steve Weinberger.

For the past three years I have been fortunate to be employed teaching English and English linguistics in the Institute of Language and Culture Studies at Hokkaido University in Sapporo, Japan. My experience there has provided me with the opportunity to collect a substantial portion of the data reported in this dissertation as well as the chance to view second language acquisition (or lack of it, in many cases) first hand. Down in the trenches, as it were, I have gained a new appreciation for the quest of second language teachers looking to linguistic theory and second language theory for answers to acquisition puzzles as well as guidance in teaching methodology. This experience has also made me recognize, perhaps more than I wanted to, that second language research does not always have, nor should have, direct application to the language classroom.

I would like to thank several members of the English department, in particular, who either directly or indirectly contributed to the completion of work: Katsunosuke Namita, the dean of the Institute, who took interest in my research and encouraged me more than once with a "Gambatte kudasai!"; Masanobu Ueda, a fellow linguist, whose stimulating conversations always made me feel less isolated from the linguistics community; and Yujin Yaguchi, another combination teacher/doctoral student, who, though not a linguist, could certainly empathize with the challenges of trying to complete a dissertation long-distance while teaching full-time. I would also like to thank my native-speaker colleagues – Christopher Glick, Mark Holst, Jennifer Morris, Paul Stapelton and Joseph Tomei – who always encouraged me and never complained when dissertation research prevented my full participation in our joint teaching projects.

Of course, I never would have gone to graduate school in linguistics had it not been for my experience in the linguistics program at Brandeis University. I had the good fortune to be an undergraduate there at a time when the linguistics program was led by a handful of remarkable scholars. Though I took "Intro to Linguistics" only because I had to fulfill a social sciences requirement (and Psych 101 was already filled), once Jane Grimshaw tapped us into an awareness our own implicit linguistic knowledge I was hooked; since then she has been a model of teaching that I continually strive to approximate. I would also like to express my appreciation to the other members of the department at that time, Ray Jackendoff, Joan Maling, Alan Prince and Edgar Zurif, for creating such a stimulating intellectual environment; thanks also to Arild Hestvik, Sarah Rosen and Jacqueline Toribio for guiding me through countless problem sets, as teaching assistants. But the person I owe the most gratitude to, for stirring within me a desire to do linguistic research and for showing me, by her own example, that a woman could be both a distinguished scholar and devoted wife and mother, is Moira Yip. As my mentor, she not only taught me to do phonological research, she allowed me a glimpse into the personal and professional life of an academic, giving me a taste of a vocation that, in large part due to her wonderful example, I have gone on to pursue.

This work would not have been possible without the several sources of financial support I have received throughout my time at McGill: Eileen Peters McGill Major Fellowship, Friends of McGill Major Fellowship, Arthur C. Tagge McGill Major Fellowship, and Social Sciences and Humanities Research Council Fellowship number 410-89-0515 to Glyne Piggott, number 410-92-0047 to Lydia White, and FCAR team grant number 91-ER-0578 to Lisa Travis, Lydia White and Mark Baker.

Finally, I would like to thank those who have had the most profound influence on me, my parents and siblings. My parents, Brenda and Roland, and my brother and sister, Robert and Lori, have given me more love than I could possibly need and have provided unwavering support for everything I have done (well, maybe not the hot pink hair phase), believing in me at many times when I, myself, did not. The one thing they did *not* do, and for this I think any graduate student would be grateful, is ask "So, *when* exactly are you going to finish?" too many times. Even though I have never been able to adequately explain to them what it is that linguists do, it is enough for them that I am doing something I love. They truly are "the wind beneath my wings".

Little did I know when I entered McGill's graduate program that I would meet the man of my dreams. Fortunately for me he was also studying linguistics and we have been a team, in linguistics and love, ever since. My husband, John Matthews, has touched this research and my soul more than any other person. He has generously contributed his time and ideas to the advancement of my work, never once complaining. His intellect is on every page of this work and his unconditional love is in every fiber of my being. John, I only have to look at you to know that God's love is real and that I have been blessed from above. *Enfin mon égal.*

This dissertation is dedicated to the memory of my father, Roland Brown, who always encouraged me to reach higher than I could comfortably go and instilled in me a deep desire to know "why", which ultimately led me to pursue and complete a Ph.D. Though he did not live to see me graduate from university, I am confident that he is now beaming with pride, elbowing the nearest angel to whisper "Yep, she's my daughter, I knew she could do it!" Dad....I did it.

Thank you.

Acknowledgements for the Papers

As each of the chapters that comprise this dissertation is an independent paper, individual acknowledgements for each are necessary; these are included together here, along with information regarding the publication status of each paper.

Chapter 1:

"The Feature Geometry of Lateral Approximants and Lateral Fricatives" appears in Harry van der Hulst and Jeroen van de Weijer, editors, (1995) *Leiden at Last: Proceedings of the 1st annual HIL Phonology Conference*, published by the University of Leiden, Holland. An earlier version also appears under the title "On the representation of laterality" in John Matthews, editor, (1993) *McGill Working Papers in Linguistics* 8.2.

> Acknowledgements: Many of the ideas presented in this paper originated from discussions with John Matthews; the quality of this research has greatly improved as a result of his generous contributions. I am also indebted to Glyne Piggot for encouraging me to tackle the issue of laterality as well for his own scholarship in this area, which was to form the foundation of my own investigations. Valuable comments were also provided by Heather Goad, Phil Hamilton, Harry van der Hulst, Takako Kawasaki, Jaye Padgett, Joe Pater, Keren Rice, Sharon Rose and an anonymous reviewer. I gratefully acknowledge support for this research received from a Friends of McGill Major Fellowship to the author, and SSHRC research grant no. 410-89-0515 to Glyne Piggott.

Chapter 2:

"The Role of Feature Geometry in the Acquisition of Phonemic Contrasts" will appear in Martha Young-Scholten and S.J. Hannahs, editors, *Focus on Phonological Acquisition*, published by John Benjamins Publishing Company. Portions of it also appear under the title "The acquisition of segmental structure" in John Matthews and Lydia White, editors, (1993) *McGill Working Papers in Linguistics* 9: *Special Issue on Language Acquisition.*

Acknowledgements: This research was first presented at the 17th annual Boston University Conference on Language Development, October, 1992, where useful questions and comments were received from Jane Grimshaw and Alec Marantz. Helpful comments were also received on a written version from Catherine Best, Jane Fee, Paula Fikkert, Heather Goad, Claartje Levelt, Joe Pater, Glyne Piggott, Keren Rice and Lydia White. I would like to especially acknowledge and thank John Matthews, my co-author on this paper, for his help during the collection of experimental data, initial presentation of our findings, and assistance throughout the preparation of this manuscript for publication. I gratefully acknowledge the participation of the teachers, parents and children at the McGill Community Daycare Centre. In particular, I would like to thank Lise Simard for her support throughout the long process of the data collection. Support for this research was provided by an Arthur C. Tagge McGill Major Fellowship to the author, FCAR team grant no. 91-ER-0578 to Lisa Travis, Lydia White and Mark Baker, and by SSHRC no. 410-89-0515 to Glyne Piggott.

Chapter 3:

"The Role of the L1 Grammar in the L2 Acquisition of Segmental Structure" is currently under revision for *Second Language Research.* An earlier version appears under the same title in John Matthews and Lydia White, editors, (1993) *McGill Working Papers in Linguistics 9: Special Issue of Language Acquisition.*

Acknowledgements: I would like to thank the audiences at the 18th annual Boston University Conference on Language Development (January, 1994) and the 1994 Second Language Research Forum for their insightful questions and remarks; comments from Barbara Hancin-Bhatt, Lise Menn and Barbara Pearson were especially relevant. John Matthews deserves my heartfelt thanks for many discussions on several aspects of this paper and for his help in preparing the stimuli and compiling the data. Instructive comments were also received from members of the Acquisition of Phonology research group at McGill University, in particular, Heather Goad and Lydia White. This research would not have been possible without the generous assistance of Dongdong Chen and many of the Japanese graduate students at McGill University, especially Takako Kawasaki and Makiko Hirakawa. I am grateful for the financial support provided by an Eileen Peters McGill Major Fellowship to the author, and SSHRC research grant no. 410920047 to Lydia White during the completion of this research.

Chapter 4:

"The interrelation between speech perception and phonological acquisition from infant to adult" is under revision to appear in John Archibald, editor, *Second Language Grammars*, published by Blackwell Publishers. Portions of it also appear under the title "English tongues, Japanese ears: A theory of phonological interference", in *Language & Culture: Journal of the Institute of Language & Culture Studies* 30, Hokkaido University, Sapporo, Japan.

Acknowledgements: I would first like to thank John Archibald for the opportunity to bring a substantial portion of my theoretical and experimental L2 research together in a (cohesive?) format. Several people have contributed to the completion of this research: I am grateful to Joseph Tomei and John Matthews for lending their voices for experimental stimuli, to Masanobu Ueda for his assistance with the DAT equipment at Hokkaido University, and to Michie Namita for recruiting Korean and Chinese subjects in Sapporo, Japan. I would also like to thank Heather Goad and Lydia White for their guidance throughout this research, which often occurred long-distance via email. Finally, this research would be significantly diminished without the generous contribution of John Matthews to the design and presentation of this study. This research was supported by SSHRC research grant #410920047 to Lydia White.

Thesis guidelines require me to make explicit the extent to which others have contributed to this dissertation. This is only an issue for Chapter 2, which is co-authored by John Matthews. The study that comprises Chapter 2 is truly a collaborative work, such that the division of labor cannot be easily distinguished. The design of the study, preparation of experimental materials and collection of data in the daycare was carried out equally and collectively by both authors. Preparation of the manuscript, including the initial writing and subsequent revisions, was also completed jointly.

Thesis Guidelines for Manuscript-based Dissertations

Each of the chapters of this dissertation consists of an independent paper that has been submitted for publication (see acknowledgements for the papers). In accordance with the guidelines of the Faculty of Graduate Studies, the following indented paragraphs are included to inform the reader of Faculty regulations concerning the manuscript-based dissertation format:

> Candidates have the option of including, as part of the thesis, the text of one or more papers submitted or to be submitted for publication, or the clearly-duplicated text of one or more published papers. These texts must be bound as an integral part of the thesis.
>
> If this option is chosen, connecting texts that provide logical bridges between the different papers are mandatory. The thesis must be written in such a way that it is more than a mere collection of manuscripts; in other words, results of a series of papers must be integrated.
>
> The thesis must still conform to all other requirements of the "Guidelines for Thesis Preparation". The thesis must include: a table of contents, an abstract in English and French, an introduction which clearly states the rationale and objectives of the study, a review of the literature, a final conclusion and summary, and a thorough bibliography or reference list.
>
> Additional material must be provided where appropriate (e.g., in the appendices) and in sufficient detail to allow a clear and precise judgement to be made of the importance and originality reported in the thesis.
>
> In the case of manuscripts co-authored by the candidate and other, the candidate is required to make an explicit statement in the thesis as to who contributed to such work and to what extent. Supervisors must attest to the accuracy of such statements at the doctoral oral defense. Since the task of the examiners is made more difficult in these cases, it is in the candidate's interest to make perfectly clear the responsibilities of all of the authors of the co-authored papers.

Introduction to the Thesis

Within standard generative linguistics, a speaker's phonological knowledge is believed to comprise a set of abstract representations, a set of well-formedness constraints and a set of rules that govern the realization of those abstract forms. Accordingly, one of the primary goals of phonological theory within this framework is to determine the precise content and structure of phonological representations in a speaker's mental grammar and to formulate rules which account for the relationship that holds between these representations across different levels of the phonology (e.g., underlying and surface forms). Acquisition researchers and psycholinguists working in this paradigm have, at the same time, endeavored to demonstrate that the constructs posited by theoreticians are not merely formal tools for describing a linguistic system, but correspond to psychologically real entities that are learnable and have psycholinguistic ramifications once acquired.

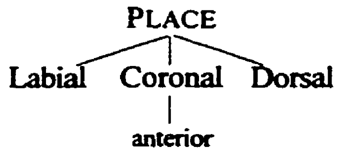
With respect to knowledge of phonemes, the introduction of the theory of segmental representation known as Feature Geometry (e.g., Clements, 1985; Sagey, 1986) has provided significant insights into both the nature of segmental representations and the processes which act upon those representations. As a result of advances in this framework, the grammar has come to be viewed as containing highly articulated representations and a small set of simple phonological rules whose operation and particular effects follow directly from the content and structure of segmental representations.[1] Though long recognized that segments are themselves comprised of

[1] There has been a recent shift away from conceiving of the grammar as a set of rules and constraints to viewing it as containing only a set of universal constraints that govern the well-formedness of output (surface) forms, as put forth in Optimality Theory (McCarthy & Prince, 1993 and Prince & Smolensky, 1993; the role of constraints is similarly emphasized in Harmonic Phonology, Goldsmith, 1993, and in the TCRS framework, Paradis & LaCharité, 1993). At the same time, several researchers working within Optimality Theory have downplayed the role of segmental representations, including the hierarchical organization of features within the segment, placing the burden of accounting for segmental processes entirely onto the constraints and their complex interaction (see e.g., Cole &

subcomponents, called distinctive features (e.g. Chomsky & Halle, 1968; Jakobson, 1941; Trubetzkoy, 1939), the innovation that these features are organized within the segment in hierarchical arrays (i.e., not unordered bundles) has enabled phonologists to explicate the range of segments that languages use to construct their phonological systems and to explain the range of phonological processes attested cross-linguistically, including the behavior of individual segments in those processes.

The hierarchical relations encoded in feature geometric representations – dependency and constituency – capture structurally many of the cross-linguistic observations that heretofore were merely stipulations in the theory. Dependency relationships between features, as shown in (1), explicate the fact that not all possible feature combinations yield sounds that are attested cross-linguistically (note: for expository purposes, only a portion of the Feature Geometry has been given).

(1)



We see here that the feature [anterior] is a dependent of the feature [coronal]: the presence of [anterior] in a representation entails the presence of [coronal][2] and, conversely, [anterior] may only attach to [coronal]. This unique structural relation between these two features explains why the anterior sounds that define a natural class are invariably coronal and why phonological processes that target [coronal] will affect both anterior and non-anterior coronal segments.[3]

In addition to dependency relations, the features that define a particular class of

---

Kisseberth, 1994; Padgett, 1994; Pulleyblank, 1994). The arguments presented in this thesis will be couched in terms of feature-geometric representation and the exploration of first and second language acquisition will demonstrate that there are significant theoretical and empirical advantages to maintaining an internal segmental organization. A discussion of the conceptual and empirical differences between the OT and Feature Geometry frameworks, and their implications for L1 and L2 acquisition, can be found in Appendix E.

[2] The same is true if assume we that features are bivalent (i.e., [+/– anterior]); in this case, the presence of either value of the dependent feature entails the presence of the superordinate feature.

[3] Traditionally, [anterior] was used to define labial segments and the set of coronal segments articulated at or in front of the alveolar ridge (e.g., dental and alveolar segments). However, since labials and anterior coronals do not define a natural class, [anterior] has been reinterpreted to define only a subset of coronal articulations.

sounds are organized into constituents, such as PLACE above. Since phonological processes manipulate constituents of segmental structure, the features of a constituent will pattern together in phonological operations; for example, a process that targets the PLACE node necessarily affects all places of articulation; in other words, no process can target both labial and coronal segments, to the exclusion of dorsal segments, because the features [labial] and [coronal] do not form a constituent in the hierarchy. Thus, the fact that phonological processes invariably target natural classes of sounds is also formally captured by the relations encoded in the feature hierarchy.

Feature Geometry has, therefore, not only improved our conception and theoretical characterization of phonological representations and rules, thereby resulting in a more tightly constrained theory, it has also made it possible to provide a unified analysis of heretofore seemingly disparate phenomena. Processes such as place-of-articulation assimilation, whereby a nasal segment, for example, takes on the place of articulation of a following stop consonant (regardless of the stop's place of articulation) are widely attested cross-linguistically, yet in linear frameworks, such as The Sound Pattern of English (Chomsky & Halle, 1968), these very common processes were expressed as three separate rules involving [anterior] and [Coronal]; in Feature Geometry theory such assimilation is viewed as a single operation that targets the superordinate PLACE node, thus capturing the elements of the segments these processes affect as well as the prevalence of these processes. Furthermore, and perhaps more importantly, this theory provides a principled explanation for why certain phonological processes or logically possible feature combinations are NOT attested across the world's languages.

The full set of distinctive features and their hierarchical organization constitutes a universal Feature Geometry. While no one language utilizes all of the components of this geometry, every phoneme in the world's languages can be represented in terms of the features and structural relations encoded there and, conversely, the representation of every phoneme conforms to this hierarchical organization. Thus, as a universal principle of feature organization, Feature Geometry theory is hypothesized to be a component of the biologically endowed language faculty, Universal Grammar (UG).

Couched within the larger generative paradigm, Feature Geometry is intended to be a characterization of the mental construct underlying speakers' knowledge of phonemes, rather than merely a formal tool for describing a phonological system. Not surprisingly then, in addition to cross-linguistic phonological research, there has been considerable research dedicated to uncovering the acoustic and articulatory foundations of the content and hierarchical structure of segmental representations (see e.g., Browman & Goldstein, 1986, 1989; Halle, 1983; Stevens, 1989). The feature organization, which effectively represents the vocal tract, provides a more direct relationship between the phonological representation and its phonetic interpretation, elucidating the relationship between the abstract linguistic system and its realization in the physical world; the theory of Feature Geometry enables us to structurally capture the connection between what is a possible speech sound (i.e., its acoustic and articulatory properties) and its phonological behavior in the grammar. This convergence of the phonetic properties of segments and their abstract phonological behavior, while not requisite for the generative position, lends support to the claim that the hierarchical organization of speech sounds is a basic property of the human language faculty.

The aim of this dissertation is to examine the acquisition of feature geometric representations in first and second language acquisition in order to provide evidence for the claim that the theory of Feature Geometry is a property of Universal Grammar and that feature geometric representations are, in fact, mental constructs in a speaker's grammar and, as such, underlie language use. By investigating the representation of /l/ and /r/ and acquisition of these segments, as well as others, I demonstrate that the theory of Feature Geometry actively guides and constrains the acquisition process and that once acquired, feature geometric representations have psycholinguistic consequences for the perception of speech sounds. Below, I briefly introduce the four independent papers that comprise this dissertation and indicate how the results from each one contribute to the goal of providing support for the claim that the hierarchical nature of segmental representations is a property of UG and that segmental representations correspond to psychological properties of the human mind.

Although the theory of Feature Geometry is by now well developed and empirically supported by cross-linguistic data, there is still some disagreement in the literature as to the exact arrangement of specific features and the corresponding representation of segments containing those features (see e.g., Clements & Hume, 1994; Rice & Avery, 1991b; McCarthy, 1988). The ongoing task of research in this area, then, is to determine conclusively which phonological features should be recognized and where they are to be positioned in the universal hierarchy of features. Chapter 1 of this dissertation, "The feature geometry of lateral approximants and lateral fricatives", contributes to this task by considering the representation of laterality and the status of the feature [lateral] in the universal hierarchy. Lateral segments have traditionally been analyzed as being distinguished from other segments by the feature [lateral]. However, there is considerable controversy regarding the proper positioning of [lateral] in the hierarchy and, therefore, the correct representation of lateral segments (see e.g., Levin, 1988; Rice & Avery, 1991b; Shaw, 1991; Yip, 1990). In addition, existing proposals for the representation of laterality cannot account for the full range of behavior of these segments cross-linguistically: positing [lateral] as a dependent of the SONORANT VOICE node fails to explain why, in some cases, laterals are created during place assimilation, and positing [lateral] as a Coronal dependent incorrectly predicts that place assimilation will produce lateral segments cross-linguistically and that processes involving sonorancy will NOT affect laterals. Moreover, these proposals are based entirely on the behavior of sonorant laterals, effectively ignoring any insights into laterality that examination of non-sonorant laterals could potentially provide.

The evidence presented in this chapter, from the behavior of both sonorant and non-sonorant laterals in a variety of phonological processes, indicates that laterality is not a singular property, but is crucially contingent upon properties of both place of articulation and degree of articulation, thus resisting standard treatment in feature geometric terms. On theoretical grounds, I argue that [lateral] is not a phonological feature, but rather that laterality is actually a phonetic property of segments that derives from the implementation of a particular phonological structure.[4] Viewing laterality in this way enables us to account for the fact that lateral segments are invariably coronal yet do not participate cross-linguistically in phonological processes that target coronal segments.

Eliminating the feature [lateral] from the universal hierarchy of phonological features, however, has consequences for several aspects of the phonological system. Most importantly, it requires that lateral segments be distinguished from other approximant (or fricative) segments in the speaker's mental grammar by some means other than that feature. Thus, in addition to considering the representation of laterals, I discuss the representation of the non-lateral approximant /r/ and the non-lateral fricative /s/, and suggest that both are distinguished from their lateral counterparts by the feature [coronal]. Removing [lateral] from phonological representations also predicts that this feature will not be a factor in the acquisition of phonemes: it should play no role in predicting the order of phoneme acquisition and phonological rules posited by children should not make crucial reference to this feature. Rather, the feature [coronal] is predicted to be intimately linked to the acquisition of /l/ and /r/ (and /s/ in those languages that contain a lateral fricative) and to phonological processes which affect these two segments. Finally, excluding [lateral] from the universal feature hierarchy (and, therefore, from representations posited in a speaker's grammar) entails that the psychological reflexes of feature-geometric representations (whatever they may be) will not stem from the laterality of a segment, but rather from its coronality. While the phonological representations proposed for /l/ and /r/ are supported by cross-linguistic data, assessing the predictions that these representations make for the acquisition and use of this knowledge requires us to look outside the phonological system proper, to its development in first and second language learners.

The second paper of this dissertation, "The role of feature geometry in the acquisition of phonemic contrasts", explores the consequences for first language acquisition of postulating that /l/ and /r/ are distinguished by the feature [coronal], while

[4] Spencer (1984) also proposes eliminating the feature [lateral]; his proposal is discussed in Chapter 1, footnote 1.

at the same time investigating the role of UG in the acquisition of phonological representations, more generally. As part of the phonological knowledge contained in the biologically endowed language faculty, UG, the universal feature hierarchy (including its content and the dependency and constituency relations encoded there) should provide the child information about what phonemic contrasts are possible in the world's languages and guide the acquisition of those feature-geometric representations. Researchers have, in fact, demonstrated that children's phonological performance can be analyzed in terms of feature-geometric representations and rules that operate on those hierarchical representations (see. e.g., Dinnsen, 1993; Goad, 1996; Ingram, 1989; Levelt, 1994; Spencer, 1986, *inter alia*). However, to date it has not been established precisely how those representations are acquired in the first place; similarly, it has not yet been investigated whether Feature Geometry theory plays a discernible role in the acquisition process.

Chapter 2 develops and experimentally tests a theory of phoneme acquisition whose central claim is that the universal feature hierarchy encoded in UG actively constrains the acquisition of segmental representations by providing the material from which representations will be constructed and determining the order in which they will be acquired. It is hypothesized that when the child detects that two sounds are used contrastively in the ambient language, the structure that distinguishes those two segments is added to the emerging Feature Geometry in the child's mental grammar (Rice & Avery, 1991a, 1995). The experimental data presented here show that segmental representations are indeed gradually built up, in a systematic order, until all of the target language phonemes are differentiated, and that /l/ and /r/ in particular are only distinguished once [coronal] has been added to the child's grammar. By establishing how children acquire segmental representations, this research demonstrates that feature-geometric representations are indeed acquired on the basis of positive evidence and that knowledge of the hierarchical arrangement of features is a component of UG, thereby providing indirect support for the proposed representations of /l/ and /r/, and for the theory of Feature Geometry itself.

Still at issue, though, is whether feature-geometric representations have any psychological validity; can we find evidence that these representations underlie language use? If phoneme representations are indeed mental constructs in a speaker's mind, then, once acquired, we might expect psycholinguistic effects that reflect the nature of those representations. For that evidence we turn to mature speakers. The results of chapter 1 (i.e., that /l/ and /r/ are distinguished by the feature [coronal]) combined with the results of chapter 2 (i.e., that segmental representations are constructed on the basis of the learner's detection of contrastive use of those segments in the input) makes some interesting (and testable) predictions for how the native grammar will affect second language acquisition. The third paper presented in this thesis, "The role of the L1 grammar in the L2 acquisition of segmental structure", synthesizes the findings from the first two papers in order to develop a theory of how the L1 grammar, in particular, the mature feature geometry, influences the course and relative success of acquiring a second phonological system. Most second language acquisition researchers readily acknowledge that the L1 grammar exerts a tremendous influence on L2 acquisition, though they are less certain as to which aspects of the grammar are responsible for this interference and precisely how it arises (see e.g., papers in Schwartz & Sprouse, 1996). This paper represents the first attempt to construct a theory of L1 interference which integrates the findings of infant speech perception research and L1 phonological acquisition and, in doing so, suggests an origin of the L1 phonological system's effect on the speech perception system and the mechanism by which it subsequently constrains L2 acquisition.

Examination of the development of speech perception, in conjunction with L1 phoneme acquisition, indicates that a child's perceptual capacities worsen as his or her ability to discriminate phonemes structurally in the grammar improves. This is taken as evidence that the elaboration of Feature Geometry in the child's grammar imposes the boundaries according to which phonemes will be categorized perceptually. Thus, not only does the acquired hierarchy of features reflect which segments are contrastive in the child's grammar, it also corresponds to the organization of the speech perception system. This proposal, of course, has important implications for L2 acquisition. Since accurate perception of contrastive segments triggers acquisition of the structure that differentiates those segments, learners should only acquire those non-native phonemes which they detect in the input. These predictions are tested by investigating the acquisition of the /l–r/ contrast by native speakers of Chinese and Japanese. The experimental results indicate that L2 learners will, in fact, accurately perceive only those non-native contrasts that are distinguished by a feature already present in their native grammar and that they will acquire only those contrasts which they perceive as distinct. This suggests that the feature geometry elaborated in the child's mental grammar in first language acquisition mediates the perceptual system, defining the boundaries of their perceptual capacities. Thus, the research presented here establishes that feature-geometric representations have psychological reflexes in the perception of speech sounds, strongly supporting the view that a speaker's knowledge of phonemes is hierarchical in nature and truly a property of the mind.

The final paper presented in this thesis, "The interrelation between speech perception and phonological acquisition from infant to adult", further develops the theory of L1 influence outlined in the previous chapter and extends this research by exploring the acquisition of several English contrasts by native speakers of Japanese, Chinese and Korean. Comparison of these three language groups enables us to determine more conclusively whether it is the *phonological* features or the *phonetic* features of the L1 that guide speech perception. The differential success that these speakers have in acquiring a given non-native contrast, as well as the differential success that speakers of a particular L1 have in acquiring different contrasts, is shown to follow directly from the (phonological) featural properties of their L1 grammars. By isolating the properties of the L1 grammar that impinge upon L2 acquisition, we firmly establish that what appears to be "partial influence" of the L1 grammar is, in fact, total influence of the L1 phonological system on the L2 acquisition process: facilitating proper perception of the input and, hence, successful acquisition in some instances and impeding accurate perception and, therefore, preventing acquisition in others. How this effect of the native grammar changes over the course of L2 development is also addressed experimentally by comparing the perceptual capacities and phonological development of beginners and more advanced learners. Thus, the theory of L1 influence and L2 segmental acquisition presented here, which relies heavily on the representations proposed in Chapter 1 and the theory of phonological acquisition espoused in Chapter 2, accurately accounts for cross-linguistic, as well as developmental, patterns of L2 phoneme acquisition.

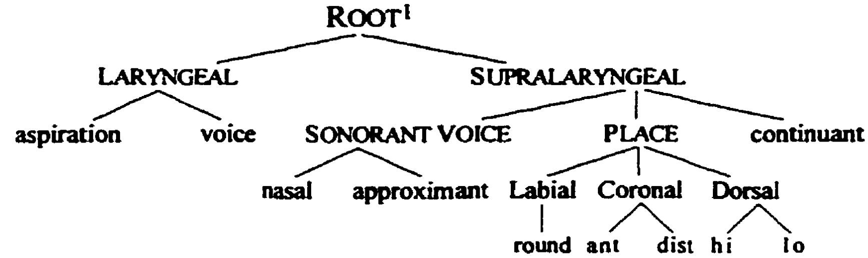
The dissertation concludes with a discussion of the contributions that the combined research presented here makes to our understanding of what constitutes knowledge of phonemic contrasts, how this knowledge is acquired and, once acquired, how this knowledge is used in the processing of speech sounds. This discussion is extended in Appendix E with an examination and evaluation of the implications of a recent shift in phonological theory for the theories of first and second language phoneme acquisition developed in this dissertation.

Preface to Chapter 1

In order to explore the acquisition of segmental representations by children and examine the psycholinguistic effects of these representations in mature speakers, we must first establish how to best formally characterize a speaker's knowledge of the phonemes used in his or her language. Only once we have determined the form and content of these representations can we consider how that knowledge is acquired and subsequently used in language comprehension and production. Thus, Chapter 1 will explore the representation of lateral segments, and other closely related segments, in order to provide the theoretical underpinning for our subsequent exploration of the acquisition and use of these segmental representations in first and second language acquisition. Before we examine the representation and phonological behavior of lateral segments, though, it will be necessary to briefly review a few of the assumptions underlying the theory of segmental representation that will be adopted in this, and following, chapters.

The theory of segmental representation that will be pursued here is known as Feature Geometry (Clements, 1985; Sagey, 1986). The universal set of features and their hierarchical organization can be illustrated in a model of Feature Geometry, such as the one presented in (1).

(1)



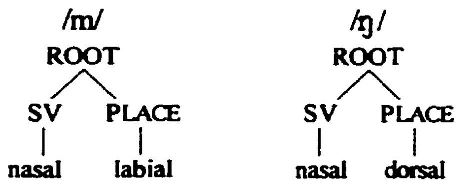
---

[1] The abbreviated features [ant], [dist], [hi] and [lo] stand for [anterior], [distributed], [high] and [low], respectively.

Thus, Feature Geometry differs from its predecessors in postulating that the distinctive features which comprise segments are arranged within the segment into hierarchical constellations (i.e., feature geometries), rather than consisting simply of unordered bundles (e.g., *Sound Pattern of English*, Chomsky & Halle, 1968; henceforth, SPE). Each of the features in this geometry exists on its own tier and can, thus, function in the phonological system of a language independently; note, however, that these autosegmental features are also organized into groups of features (i.e., constituents), which are dominated by a single feature or "node". Thus, although the different features cannot "see" each other (because they are organized on separate tiers), they may function as a constituent in phonological processes because at a higher level of organization they constitute a single unit.

While no one language manipulates all components of this universal Feature Geometry, every phoneme in the world's languages can be represented in terms of the features and structural relations present in this geometry. Phonemes are distinguished from one another in a grammar by their unique structural representation (i.e., feature geometry). The representation of a given segment will be a subset of the entire universal geometry above; its precise representation will depend on which segments it contrasts with in the particular inventory, yet it will always conform to the hierarchical relations encoded in (1). An example of two such feature-geometric representations is given in (2); note that SV stands for Sonorant Voice, which will be explained below (also, the superordinate structure that is not relevant to this discussion has been omitted).

(2)



Both of these segments contain the feature [nasal], indicating that they are articulated with a flow of air through the nasal cavity; however they contrast in their specification for place of articulation: the geometry for /m/ contains the feature [labial], while /ŋ/

contains the feature [dorsal]. Any language in which /m/ and /ŋ/ are phonemes will contain these representations in its grammar.

Though Feature Geometry theory has rejected many of the assumptions of linear phonology, it still recognizes features to be the primitive components of phonological representation. Thus, an important issue in this framework, and an ongoing task for theoreticians, concerns which features should be formally recognized. Historically, features have been defined such that each one corresponds to both an articulatory and an acoustic property of sounds (Jakobson, Fant & Halle, 1952). However, within the SPE framework (Chomsky & Halle, 1968), there was a shift towards defining (most) distinctive features in terms of articulatory properties, and that trend has continued into non-linear phonology (see especially Sagey, 1986).

While Feature Geometry theory has basically adopted the set of SPE features, it has also introduced some new ones as well. Since Feature Geometry organizes features into hierarchical constituents, a distinction is made between features (i.e., terminal features) and nodes (i.e., the structure which organizes those features); the nodes, themselves, are further divided into two types: Content nodes, which often correspond to traditional features, and Organizing nodes (also referred to as class nodes by Clements, 1985), which organize Content nodes and/or features. Finally, the ROOT node organizes all Organizing nodes. Looking back at the geometry in (1), we see, for example, that PLACE is an organizing node, since it dominates the nodes that define place of articulation, while Coronal is a content node, and [anterior] is a terminal feature. Though there is a three-way distinction between Organizing nodes, Content nodes, and features, for ease of exposition, I will typically refer only to nodes and features, treating content nodes, such as coronal, as features.

Despite certain differences between organizing nodes and features, the evidence used to determine whether a feature or node should be recognized in the universal Feature Geometry is the same: in order to be recognized in the hierarchy, a feature or node must be shown to be manipulated, either as a trigger or a target, in some phonological process. When a phonological rule makes unique reference to a given phonological feature, the phonologist has evidence that that feature is indeed a member of the universal hierarchy; similarly, when two or more features regularly pattern together in phonological processes, the phonologist has evidence that those features form a constituent that is headed by a unique Organizing node. We will see in this chapter that there is still no consensus in the literature as to the universal set of features.

Closely related to the issue of which features to recognize is the question of whether features are bivalent (i.e., being specified for both (+) and (–) values) or monovalent. Here, again, evidence for a particular value of a feature comes from its demonstrated involvement in phonological processes. Traditionally, each feature has been assumed to have both a positive and negative value; however, several aspects of Feature Geometry theory led to the adoption of monovalent features. The introduction of organizing nodes, which were simply present in or absent from a representation, opened up the possibility that the distinction between (+) and (–) values of a feature might be captured by the mere presence or absence of that feature in a representation. Furthermore, the move toward an articulator based theory in which certain nodes, such as Coronal, represented the articulation itself suggested that not all features were bivalent: the presence of [–coronal] in a representation, which would entail the lack of this articulator's involvement, is incongruent with the presence (in the same representation) of its dependent, [+anterior], which would designate the active involvement of the coronal articulator. Thus, it was postulated that the Organizing and Content nodes, such as PLACE and Coronal, were monovalent, but that the terminal features, such as [anterior], were bivalent (Sagey, 1986). Mounting phonological evidence, however, suggests that only one value of most features is ever crucially referred to in the formulation of phonological rules (see e.g., Steriade, 1987, on [round]; Itô & Mester, 1989, on [voice]); considerations of parsimony (i.e., fewer options available to the grammar), then, favor monovalency. Therefore, I will assume here that all phonological features are monovalent and it is the mere presence of a feature in the representation of a segment that designates the active involvement of the relevant articulator, while the absence of a feature from a representation entails that the corresponding articulator is not active in the production of that segment (Anderson & Ewen, 1987; Avery & Rice, 1989; van der Hulst, 1989).

A separate yet related issue is underspecification. In addition to determining which features to include in the universal Feature Geometry, and their respective values, phonologists must also wrangle with the issue of how many features to include in any given representation. Much of the work in phonological theory, as in linguistic theory in general, has been guided by the presumption that underlying representations contain only idiosyncratic properties; any redundant, or predictable, information is absent from these representations and will be supplied during derivation. If we assume that features are monovalent, then some of the redundant information will be absent from segmental representations; for example, since [–continuant] is no longer a possible feature, a coronal nasal will not be specified for [–continuant], although it is, in fact, non-continuant. Yet monovalency does not remove all of the redundancy: all coronal nasals in English are anterior (i.e., there is no contrast between anterior and non-anterior nasals), thus the feature [anterior] is redundant for nasals in the English system. Eliminating redundant features from segmental representations captures formally the fact that these features are inert in the phonological system; the question is how to determine which features are redundant. Once again, involvement (or, in this case, non-involvement) of features in phonological processes or phoneme inventories provides evidence for their specification (or underspecification) in segmental representations.

Although the issue of underspecification has been the subject of much debate and has given rise to many different proposals in the literature, it continues to play a significant role in theories of non-linear phonology (e.g., Radical Underspecification: Archangeli, 1984, 1988; Pulleyblank, 1986; Minimally Contrastive Specification: Avery & Rice, 1989; Degree-2 Specification: Clements, 1988b; Contrastive Specification: Steriade, 1987). In this dissertation, I adopt Avery & Rice's theory of Minimally Contrastive Specification (MCS). This theory of underspecification, which assumes monovalent features, is inventory driven: which features a segment is specified for depends *entirely* on what it contrasts with in the particular inventory. Segments are only specified for those features which are necessary to distinguish them from other phonemes in the language, and features are introduced into a grammar only as they are required to differentiate segments in the inventory. Referring back to our English coronal nasal again, under MSC, this segment's representation will not contain the feature [anterior] because there is no contrast between anterior coronal and non-anterior coronal nasals in English; the anteriority of this coronal nasal is predictable and will, therefore, be supplied during phonetic implementation.

Thus, the issues of monovalency and underspecification refine our theory of segmental representation: eliminating redundancy (i.e., two values of features and predictable features) from underlying representations entails that these features will be inert in the phonological system. One consequence of these assumptions is that phonologists can use the involvement (or non-involvement) of particular features as direct evidence in determining which features to recognize and their positions in the universal Feature Geometry, as well as to ascertain the representation of individual segments that contain those features. Although many of the phonological features have been definitively located within the universal hierarchy, the correct positioning of a small handful still remains uncertain.

One such enigmatic feature is [lateral]; no fewer than four different proposals regarding its place in the geometry, and the corresponding representation of lateral segments, currently exist in the literature. Lateral segments have traditionally been analyzed as being distinguished from non-lateral segments by the feature [lateral] (e.g., Chomsky & Halle, 1968); lateral segments contain this feature (or, in a theory of bivalent features, are specified for [+lateral]) and non-lateral segments do not. With the advent of Feature Geometry theory, this analysis of lateral segments was maintained and phonologists set out to determine the position of [lateral] in the geometry.

This chapter uses the phonological behavior of sonorant and non-sonorant laterals in order to probe the validity of [lateral] as a phonological feature and to determine the proper feature-geometric representation of laterality. The feature-geometric representation of a segment does two things: it captures articulatory (or acoustic) properties of the segment (e.g., whether or not it is executed with the tongue blade) and it accounts for the segment's behavior in phonological processes (i.e., whether it is a trigger or target of a phonological operation, or transparent to the process). When the representation posited for a segment is correct, both the articulatory and phonological properties will fall out from that singular representation. However, lateral segments pose a problem for the theory of Feature Geometry in that the evidence from the articulatory properties of the segment and the evidence from its phonological behavior fail to converge. Lateral segments (both sonorant and non-sonorant laterals) are invariably produced with a coronal place of articulation – suggesting that [lateral] is a dependent of [coronal] – yet these segments do not always participate in phonological processes which target coronal sounds – suggesting that [lateral] is not a dependent of [coronal]. Moreover, the phonological behavior of laterals is itself ambiguous: in some instances, these segments pattern with the sonorants in a language (or non-sonorants in the case of lateral fricatives) and in other cases, they pattern with the coronals. From cross-linguistic data, it appears that laterality is dependent on properties of both place of articulation and manner of articulation.

Thus, capturing laterality in terms of feature-geometric representations is problematic: if laterality is analyzed as a property of the *place* of articulation (i.e., a dependent somewhere under the PLACE node), then we are unable to explain the fact that its realization also depends upon a particular manner of articulation. The lateral articulation is limited to the common approximant and the rarer lateral fricative; in contrast, other places of articulation, such as dorsal, combine with the full range of possible manners of articulation (e.g., velar nasals, velar approximants, velar fricatives and velar stops). On the other hand, if laterality is analyzed as a property of the *manner* of articulation (i.e., a dependent under the SONORANT VOICE node), then why is it restricted to a single place of articulation? We find, for example, coronal nasals, labial nasals and velar nasals, yet laterals are invariably coronal.

This paradox, it is suggested, can be resolved by eliminating [lateral] as a phonological feature altogether. Once we exclude [lateral] as the defining feature for lateral segments, a new type of representation is required. It is proposed that laterality is a phonetic property of certain segments that derives from their particular feature-geometric representations and evidence from a variety of phonological processes is shown to support the proposed unified representations for lateral approximants and lateral fricatives.

1 The Feature Geometry of Lateral Approximants and Lateral Fricatives

1.0 Introduction

The introduction of the theory of segmental representation known as Feature Geometry has advanced our understanding of segmental processes and provided many insights into the internal structure of segments. Many revisions have been made to the original Feature Geometry models of Clements (1985) and Sagey (1986), each one refining our conception of some aspect of the segmental geometry. However, there are still aspects of segmental representation which remain poorly understood. Perhaps the area on which we find the most debate (and least agreement) is the representation of lateral segments.

Models of feature geometry posit the feature [lateral] as a dependent of either the Root, Sonorant Voice or Coronal node. There are several reasons, however, for rejecting [lateral] as a phonological feature altogether. Empirically, the behavior of sonorant laterals is somewhat mysterious; sometimes they behave as sonorants with respect to rules that affect sonorancy, creating lateral segments where nasals otherwise result, and sometimes they behave as coronals, in rules that affect place of articulation. One of the challenges, then, to models of feature geometry is to capture this dual nature of laterals.

In addition, current claims made about the representation of lateral segments are based on the behavior of the more common (and familiar) sonorant, or approximant, lateral. In fact, no current model of feature geometry addresses the representation of non-sonorant laterals. Unfortunately, by focusing on sonorant laterals and ignoring non-sonorant laterals, we may be missing the *essence* of laterality. A definitive model of feature geometry must represent both sonorant and non-sonorant lateral segments, if we are to capture their inherent similarities. Therefore, it is imperative that in determining the representation of laterality we look at the phonological behavior of both sonorant and non-sonorant lateral segments.

The research presented here examines a wide range of segmental processes from a variety of languages (some containing sonorant laterals, some containing non-sonorant laterals and some that possess both types of laterals). Using these data, I examine four models of lateral representation that have been proposed within the theory of Feature Geometry, namely Levin (1988), Rice and Avery (1991b), Piggott (1992, 1993) and Yip (1990), and show each to be inadequate. I then offer a unified account for the representation of all types of lateral segments. In particular, I will suggest that there is no *phonological* feature [lateral], rather that laterality is a phonetic property that results from a particular phonological structure.[1]

The remainder of the paper will be organized in the following way: To begin, the behavior of lateral segments in general will be considered. I will then briefly describe four of the more prominent models of lateral representation. In §1.1, I outline my own proposal, highlighting the ways in which it differs from previous models. After reviewing some theoretical assumptions in §1.2, I will present the data that support my proposal. Evidence from coronal harmony and nasal harmony are examined in §1.3. This will be followed by a discussion of data from Navajo, Nez Perce and Welsh which shed additional light on the representation of lateral segments. Section 1.4 demonstrates how the present proposal accounts for data used to support existing models of lateral representation. In §1.5, I discuss some of the implications of my

[1] The idea that [lateral] is not a phonological feature is not new. To the best of my knowledge, Spencer (1984) was the first to suggest that the feature [lateral] should be eliminated from the universal feature inventory. However, his representation of lateral segments differs from mine. To begin, Spencer was working in an SPE-style framework where segments are defined by lists of features that have no hierarchical organization (although these representations could easily be translated into a feature geometry model). According to Spencer, laterals are distinguished from other segments by the feature [distributed]; furthermore, he does not address the representation of non-sonorant laterals nor does he discuss laterality as a phonetic feature. See Spencer (1984) for arguments to support his proposal.

proposal for the representation and acquisition of closely related segments. Finally, §1.6 concludes with a brief summary of the main claims of my proposal.

1.0.1. Summary of behavior of laterals

Before examining the existing models of lateral representation, I would like to briefly consider precisely what type of phonological behavior it is that a model must capture. One aspect of laterality that a geometry must capture is the consistent place of articulation associated with lateral segments. While phonetically velar laterals do exist in some languages (for example, Melpa, Mid-Waghi and Kanite; Ladefoged, Cochran & Disner, 1977), Levin (1988) has shown convincingly that all laterals have a coronal place of articulation underlyingly. Thus, the restriction to the coronal place of articulation must be explained.

As mentioned briefly above, laterality sometimes patterns with sonorancy. A good example of this behavior is provided by Korean. As the data in (1a–d) show, an obstruent followed by a nasal acquires the nasality of the nasal while retaining its original place of articulation. The data come from Rice & Avery (1991b:11)

(1)	a.	/kukmul/	→	[kuŋmul]	'soup'
	b.	/kakmok/	→	[kaŋmok]	'wood'
	c.	/napnita/	→	[namnita]	'to sprout'
	d.	/katʰni/	→	[kanni]	'to be the same'
	e.	/tikitliɨl/	→	[tikilliɨl]	'the letters t and l'

This process is best analyzed as leftward spreading of the Sonorant Voice node, as shown in the derivation for (1a).[2]

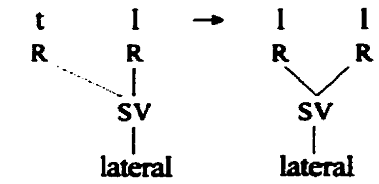
(2)
```
k    m    →    ŋ    m
R    R         R    R
 .   |          \  /
   SV            SV
   |             |
 nasal         nasal
```

---

[2] The Sonorant Voice (or SV) node is an organizing node posited in the geometry to represent sonorant properties. It has been argued for by Rice & Avery (1989) and as Spontaneous Voice by Piggott (1992, 1993).

Interestingly, the sequence /t l/ becomes [l l], as shown by example (1e). Thus, the process that spreads nasality in Korean also spreads laterality. A derivation for (1e) is given below.

(3)



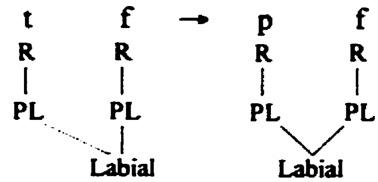
These data suggest that laterality should be represented as a dependent of the SV node. It is important to note, however, that this regressive lateral assimilation is restricted to coronal segments; if a non-coronal segment precedes the lateral, no assimilation takes place. Thus, while laterality does behave as a manner feature in Korean (and many other languages), it is usually restricted to the coronal place of articulation.

In other languages, laterality appears to behave as a place of articulation. An assimilation process from Catalan exemplifies this aspect of laterality. The data illustrating Catalan's rule of regressive place assimilation appear in (4). These forms are taken from Yip (1990), originally from Mascaró (1976).

(4)	a.	/son/	'they are'
	b.	[som pocs]	'they are few'
	c.	[soŋ grans]	'they are big'
	d.	/set/	'seven'
	e.	[sep focs]	'seven fires'
	f.	[sek cases]	'seven houses'
	g.	[sel linies]	'seven lines'

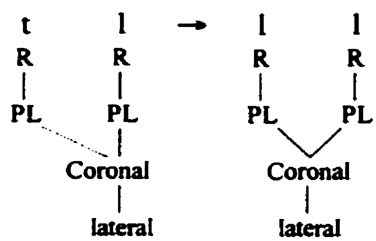
Examples (4a–c) show that an underlying alveolar nasal receives the place features of the following consonant. Examples (4d)-(4f) demonstrate that an alveolar obstruent also takes on the place features of the adjacent segment. A derivation is provided to illustrate example (4e) in which the Labial node spreads to a coronal segment, which lacks place features.

(5)



Note that, in this example, the continuancy of the /f/ does not spread to an underlying /t/, only its place features spread, creating a /p/. It is clear then that this particular rule does not affect manner features. In light of the Korean data, then, it is somewhat surprising that this process of Place assimilation also appears to create a lateral segment, as in example (4g) above. Its derivation, in which [lateral] is assumed to be a place feature is provided in (6).

(6)



Thus, there are data that suggest that laterality is a place, and not a manner, feature. This makes capturing the behavior of laterals in terms of phonological features a bit tricky. What is sometimes overlooked, though, is the fact that Catalan also has a rule of sonorant assimilation, like Korean. These data appear below.

(7) a. /set mans/ → [sem mans] 'seven hands'
b. /son liures/ → [son liures] 'they are free'

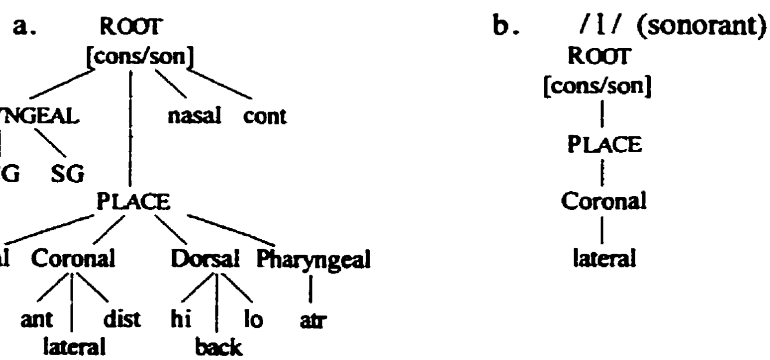
So, for instance, the coronal obstruent in the form in (7a) not only takes on the place features of the following consonant, but also its sonorancy (in this case, nasality), resulting in the form [sem mans]. Thus, the creation of a lateral segment (e.g. example (4g)) *may* actually result from the spreading of sonorancy rather than place. This hypothesis receives further support from the form in (7b). If [lateral] is a Place feature, then the sequence /n l/ should become /l l/, as a result of regressive Place assimilation.

On the contrary, though, it remains /n l/. However, if laterality spreads as the result of spreading of the SV node (as in Korean), then the surface form that we find is what we would expect; spreading of the SV node would require that the target segment be unspecified for an SV node, but because nasals do contain an SV node, spreading cannot take place and the /n l/ sequence is retained. This analysis predicts that lateral segments which are created by this rule of SV spreading should only be derived from coronal obstruents and, as far as I know, this is indeed the case. Thus, even if laterality is correctly regarded as a manner feature, the theory must explain why this manner feature is limited to the coronal place of articulation.

1.0.2. Existing Models

Now that we have examined the variable phonological nature of lateral segments, I would like to briefly describe four current feature geometry models, namely Levin (1988), Piggott (1992), Rice & Avery (1991b) and Yip (1990). For the purposes of this discussion, I will focus only on those aspects of the models that pertain to the representation of lateral segments. The first model that I will describe is espoused by Levin (1988) and is, arguably, the most widely accepted. The defining aspect of her model is that it posits the feature [lateral] as a direct dependent of the Coronal node, linking laterality directly with coronality. In this way, the model captures the observation that lateral segments always have a coronal place of articulation, underlyingly at least. Her model is shown in (8a) and the representation of a sonorant lateral is shown in (8b).

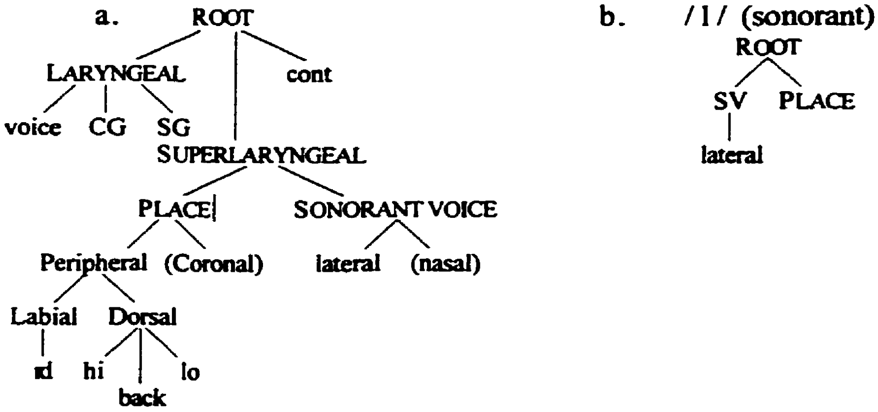
(8)



While the representation of non-sonorant lateral segments are not discussed in Levin's (1988) work, presumably they too contain the feature [lateral] as a dependent of the Coronal node; the only difference between sonorant and non-sonorant laterals, then, would be in their respective specifications for manner features.

The next feature geometry model is advanced by Rice & Avery (1991b). This model incorporates a Sonorant Voice (or SV) node. Rice & Avery consider [lateral] to be a manner feature, which is a dependent of this SV node. The Rice & Avery model is illustrated in (9), along with their representation of /l/. Note that default phonetic features are given in parentheses.

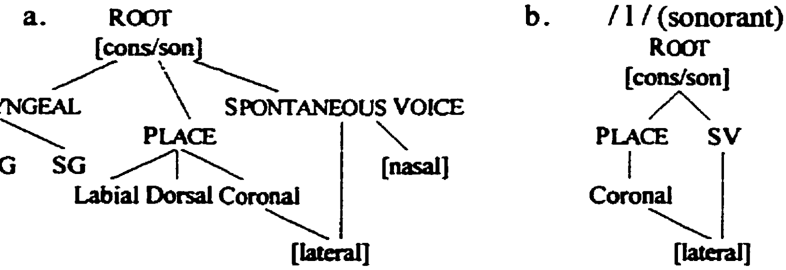
(9)



According to this model, lateral segments are not specified for Place features. Lateral segments (in fact, all segments which are specified for features under the SV node) are limited to a bare Place node by a principle called the Structure Complexity Constraint (for more on the SCC, see Rice & Avery, 1991a). The coronality of laterals results from a default rule (assumed by many theories of underspecification) that realizes a bare Place node as a phonetically coronal segment. The dependency of the feature [lateral] on the SV node is one of the strengths of the Rice & Avery model, capturing the connection between sonorancy and laterality. At the same time, though, it is also one of its weaknesses, for it limits (in principle) laterality to sonorancy, making it impossible to represent non-sonorant lateral segments with the feature [lateral]. At present, they have no proposal for the representation of fricative laterals.

The third model that will be examined is advanced by Yip (1990). Her model incorporates insights of both the Levin model and the Rice & Avery model. Yip suggests that the feature [lateral] is simultaneously dependent on both the Coronal node AND the SV node, thereby capturing the observation that the spreading of the feature [lateral] may require a target that is both coronal and sonorant. The relevant aspects of Yip's model are illustrated below.

(10)



According to this model, the feature [lateral] is solely a dependent of the SV node underlyingly, but will be linked to the Coronal node as soon as the Coronal node is introduced by the default rules.[3] Although this model does posit [lateral] as a phonological feature, it is similar in *spirit* to the model to be elaborated below as it requires a segment to be specified both for particular manner features and place features in order to be realized as a lateral segment.

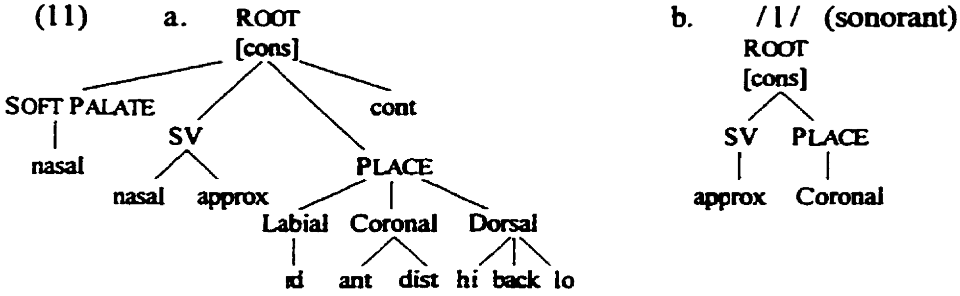
The final model that I will consider is proposed by Piggott (1992, 1993). Like Rice & Avery, Piggott incorporates a node to represent sonorancy, which he calls the Spontaneous Voice node. In contrast to the other models, Piggott maintains that laterality is not a phonological feature, but rather a phonetic property derivable from the combination of specific phonological features. In particular, he proposes that sonorant

---

[3] Yip is not specific as to exactly when, or even how, this linking is to take place. However, it is clear from the phonological behavior of laterals, that this linking must take place in the phonological component (as opposed to the phonetic component). Thus, the type of default rules that introduce the Coronal node under Yip's proposal cannot be the type of default rules discussed by Rice & Avery (1991b), which are only phonetic implementation rules.

lateral segments contain a Coronal node and the feature [approximant] (Piggott 1993).[4] His model and representation of sonorant laterals are given in (11)[5]

(11)



Unfortunately, Piggott does not propose a representation for non-sonorant lateral segments.

Each of these models attempts to capture the behavior of lateral segments across languages and each does find some empirical support. The problem, though, is that no single model can account for the entire range of cross-linguistic data. A model like Piggott's, for example, captures the interaction of laterality and sonorancy in phonological operations but does not correctly account for the behavior of laterals with respect to processes that affect place of articulation; a model like Levin's, on the other hand, captures (some of) the behavior of laterals with respect to Place assimilation, but cannot account for the behavior of laterals in processes that affect sonorancy. Moreover, none of these models currently has a means of representing non-sonorant laterals.

---

[4] The feature [approximant] was first introduced by Ladefoged (1982) and integrated into Feature Geometry Theory by Clements (1988a). It is used to represent those sounds which are characterized by close proximity, but not contact, of the articulators. Approximant sounds therefore include the liquids and the semi-vowels.

[5] Piggott's model assumes the variability of the feature [nasal]. Thus, [nasal] can be a dependent of either the Soft Palate Node or the Spontaneous Voice Node, depending on whether the particular language selects the Soft Palate as an active articulator. For arguments and data supporting this aspect of his model, see Piggott (1992, 1993). Although this aspect of his model does not play a role in the representation of lateral segments, it will be a factor when we examine data from nasal harmonies.

1.1 A Unified Model

The structures that I propose to represent lateral approximants and lateral fricatives are given in (12), (a) and (b) respectively.[6] The Air Flow (AF) node incorporated into the representation of the lateral fricative is an Organizing node, adopted from Rice & Avery (1991b) and Rice (1992), that marks stricture, distinguishing stops and continuants.

(12) a. /l/ (sonorant)

ROOT
- SPONTANEOUS VOICE
  - approximant
- PLACE

b. /ɬ/ (non-sonorant)

ROOT
- AIR FLOW
  - continuant
- PLACE

These representations incorporate many of the insights of the existing models just reviewed. In particular, I follow Rice & Avery in positing that lateral segments are not universally specified for Place features, although I do not incorporate any constraint comparable to their SCC. The idea that laterality results from a *combination* of features reflects both Yip's and Piggott's model. However, in contrast to Yip, I do not incorporate the feature [lateral] into my representations and, in contrast to Piggott, I do not believe that the Coronal node plays a crucial role in the realization of a lateral.

One of the strengths of this proposal is that it unifies the representations of sonorant and non-sonorant laterals.[7] Until now, the internal structure of non-sonorant

---

[6] The representation of lateral affricates will not be addressed in this paper. Clearly, though, an adequate theory must have a means of distinguishing them from the lateral fricatives and other coronal stops. It is clear from my proposal that lateral affricates must be distinguished from lateral fricatives and other coronal stops by some means other than Place features. The uncertain nature of the representation of lateral affricates reflects the relatively uncertain representation of affricates in general. One possible means of representing lateral affricates is as contour segments with a stop phase followed by a continuant phase. This type of representation would be sufficient to distinguish these affricates from both lateral fricatives and coronal stops, without resorting to specifications for Place features. Steriade (1993) provides a slightly different way of representing affricates, which captures the similarities between stops and affricates (see footnote 7 for a brief description of this proposal). Brown (in preparation) investigates whether Steriade's recent proposal can account for the behavior of lateral affricates.

[7] An anonymous reviewer has pointed out that in order for the proposed representations in (12) to be truly unified, the Spontaneous Voice (SV) and Air Flow (AF) nodes must be unified. While there is reason to believe that the features [approximant] and [continuant] might be represented by a single feature (to be discussed directly below), there is no comparable motivation within standard Feature Geometry theory for collapsing the SV and AF nodes. However, a recent proposal by Steriade (1993) makes it possible to further unify the representations in (12). Steriade proposes that the distinction between continuants (e.g. fricatives, approximants) and non-continuants (e.g. stops, affricates) is

laterals has been largely ignored. The representations in (12) reflect the belief that laterality is a uniform property, resulting from the specification for either [approximant] (in the case of sonorants) or [continuant] (in the case of obstruents) and a bare Place node. The lateral fricative can, of course, be specified for [voice] to obtain the contrast between voiced and voiceless lateral fricatives. The two representations are homologous in that the features [approximant] and [continuant] can be said to be functionally equivalent. This means that the presence of either one of these features in a representation signals the outward flow of air through the oral cavity. As Piggott points out (1993), within the class of sonorants, the feature [approximant] distinguishes stops (i.e. nasals) from non-stops (i.e. liquids and semi-vowels) in the same way that the feature [continuant] distinguishes stops from fricatives among obstruents. Future research must determine whether these properties are best represented by a single feature. The representations I am proposing are also identical with respect to their Place features – both are underspecified.[8]

---

structural, not featural. Very briefly, non-continuants are comprised of two articulatory phases (closure + release), represented as two aperture nodes, whereas continuants contain a single articulatory phase (release), represented with a single aperture node. Thus, fricatives and approximants, on the one hand, and stops on the other have the following representations:

continuants		non-continuants	
(approximants)	(fricatives)	(stops)	(affricates)
A_{max}	A_f	A_0A_{max}	A_0A_f

Aperture nodes (or A–positions) correspond roughly to feature-geometric Root nodes; phonological features are associated to either or both A–positions. The subscripted diacritics indicate the type of release, that is, the amount of stricture: A_0=total absence of oral airflow, A_{max}=maximal oral aperture, A_f=degree of oral aperture to produce a turbulent airstream; the type of release is not relevant for the gross distinction between continuants and non-continuants (see Steriade, 1993 for details of this proposal). An appealing aspect of this proposal is that it provides a natural means of capturing the inherent similarities between approximants and fricatives: approximants and fricatives, but not stops and affricates, contain a single A-position in their representation. This proposal would, thus, allow us to unify the representations of lateral approximants and lateral fricatives even more closely than is possible with the standard feature-geometric representation in (12).

[8] There is an interesting consequence of maintaining that laterals contain a bare Place node, rather than a Coronal node (unless required independently by contrasts in an inventory). While coronal is usually assumed to be the unmarked place of articulation, there is evidence that velar segments may be underspecified in some languages (e.g. Trigo, 1988; Rice, 1992). If a bare Place node may sometimes be realized as a coronal and sometimes as a velar, my proposed representations allow for the possibility that a lateral segment may be realized with either a coronal or a velar place of articulation. Ladefoged et al. (1977) reports velar laterals in Melpa, Kanite and Mid-Waghi. It would be interesting to investigate whether those languages that have been argued to have underspecified velar nasals also have velar laterals. Models that assume laterality to be a dependent of the Coronal node would have a difficult time accounting for these phonetically velar laterals.

Many Australian languages, which can have as many as four laterals, each with a different coronal place of articulation (e.g. dental, retroflex, etc.) would seem to challenge the claim that laterals usually lack Place features. In many cases, these various sonorant laterals do not actually contrast phonemically, but there are some languages in which they do. Because laterality is now relegated to the phonetic component, there is room for cross-linguistic variation dependent on individual inventories. That is, it may be that lateral segments in Australian languages (or any language that contrasts laterals within the coronal place of articulation) must be specified for a Coronal node and secondary features. As long as these lateral segments need not be distinguished from any other phonemes in terms of Place features alone (see §6 for more on this), there is no problem with lateral segments being specified for Place features. Quite clearly, though, as the evidence below will demonstrate, Coronal nodes are not *universally* required for the realization of a lateral segment; this is reflected in the representations in (12). This position on laterals is similar to what is assumed for most (if not all) other coronal segments: coronals are underspecified for Place features, unless required by contrasts present in an individual inventory; the phonetic realization of a coronal segment with a coronal place of articulation (e.g. /t/) is not dependent on the presence of a Coronal node in the segment's phonological representation.

Perhaps the most controversial aspect of my proposal is the claim that [lateral] is not a phonological feature.[9] Apart from the empirical evidence that I will discuss throughout the paper, there is substantial theoretical motivation for this claim. In an effort to constrain the proliferation of organizing (or class) nodes in the feature geometry, McCarthy (1988) proposes criteria for determining whether a particular node is a valid element of the universal feature geometry. According to McCarthy, each posited organizing node must meet one of the following criteria.

---

[9] What I am suggesting for laterality is what is often assumed for rhotic segments. That is, that there is no phonological feature [rhotic], yet phonologists refer to the class of r-sounds. The property of being a rhotic segment presumably results from the specification for particular manner and place features. It would seem implausible to posit a feature [rhotic], although these r-sounds clearly have something in common – in their acoustics and/or articulation. See Lindau (1985) for an interesting discussion of the various rhotics and the relationships them.

(13) An organizing node must either be
   (i) the FOCUS of an operation
   (i.e. in the case of spreading, what is being spread)
   (ii) the TARGET of an operation
   (i.e. the landing site for the spreading element)
   (iii) TRANSPARENT for an operation (i.e. not affected)
   (iv) OPAQUE for an operation (i.e. block it)

Any organizing node that does not meet one of these criteria cannot be considered a legitimate node in the geometry.

The proliferation of individual features in the geometry must also be constrained. Therefore, I believe that individual features should also be subject to the criteria in (13). What this means is that each feature must be referred to in the formulation of some phonological process, either as the focus or target of the operation, or the defining feature of a segment that is transparent or opaque to the operation. This idea is not new; similar arguments have been used to eliminate bivalent values of certain features, for example [-voice] (Mester & Itô, 1989). I would like to suggest that [lateral] fails to meet these criteria. To the best of my knowledge, there is no phonological process that makes unique reference to the feature [lateral].[10] This is not to say that lateral segments do not participate in phonological processes; I am simply suggesting that the involvement of lateral segments can be characterized in terms of some other phonological feature. In other words, when a lateral participates in a process it is by virtue of its membership in some natural class (i.e. because it is an approximant, or a sonorant or a coronal) not because of their status as a *lateral* segment. Thus, apart

[10] Steriade (1987) discusses a rule of Latin liquid dissimilation, which takes two sonorant laterals in a string and changes one of the laterals into an /r/. So, for example, /sol-alis/ becomes [sol-aris]. According to Steriade, this dissimilation fails only when the stem /l/ is separated from the suffix /l/ by an intervening /r/, as in /litor-alis/. She analyzes this process as delinking of the feature [+lateral], except when blocked by an /r/ specified for the feature [–lateral] (note that Steriade allows for bivalent features). This rule of Latin liquid dissimilation does pose a problem for the analysis proposed here and, in fact, for any model which considers [lateral] to be a monovalent features. Steriade has analyzed this process in terms of the feature [lateral] but it is entirely possible that it could be reanalyzed in terms of the feature [approximant], thereby preserving my claim that no phonological feature makes unique reference to the feature [lateral]. At this point, I do not have a good alternative analysis of these Latin facts. Let me note in passing that the reverse takes place in Georgian (i.e. a sequence of two /r/'s are dissimilated) (Fallon, 1993). For example, /gmiri–uri/ becomes [gmiri–uli]. This process receives an elegant account under my proposal: two adjacent Coronal nodes (of the /r/) create an OCP violation and one is delinked, resulting in an approximant with no Place specification (i.e. a lateral).

from the empirical evidence that we will review below, there is a strong theoretical argument for rejecting [lateral] as a phonological feature.

Now, if there is no feature [lateral], how is laterality to be represented? Empirically, it seems that lateral segments require particular specifications for manner and Place features. As we saw in Korean, the spreading of the SV node only creates a lateral segment when the target is coronal. Thus, in this way, laterality is different from other manner features, such as nasality or continuancy, which are not limited to a particular place of articulation. Maintaining that laterality is a phonetic property automatically explains why they are invariably coronal, or unspecified. No special constraint, articulatory or otherwise, is needed to ensure that lateral segments have the unmarked place or articulation – if a segment is specified for any Place features (other than Coronal, perhaps), it is simply not realized as a lateral.[11]

1.2 Theoretical Assumptions

Before we can review the relevant empirical data, the theoretical assumptions that are crucial to my analyses must be outlined. This section elaborates the assumptions that I make regarding feature values as well as the two segmental operations that I will assume in the following discussion.

1.2.1 Monovalency

An important aspect of the theory of segmental representation assumed here concerns the values for which individual features may be specified. Traditionally, features have been assumed to be bivalent, each having a (+) value and a (–) value (Chomsky & Halle, 1968). It has been observed, however, that the (–) values of particular features are never manipulated by phonological processes. Thus, with the emergence of the

---

[11] A complete theory of the representation and realization of laterality will, of course, include a theory of phonetic implementation. One of the questions that must be answered is which aspect(s) the phonological structure I am proposing cause the realization of that structure to be a lateral. In other words, what is it about the features [approximant] and [continuant] combined with a bare Place node that produces a laterally released sound? At this time, I can only speculate; additional research is required.

autosegmental framework (and subsequently, Feature Geometry), it was proposed that not all features were bivalent (see Steriade, 1987 on [round] and Mester & Itô, 1989 on [voice]). There is additional empirical and theoretical evidence that suggests that all features are monovalent (Anderson & Ewen, 1987; van der Hulst, 1989).

I assume, as do a growing number of phonologists, that features do not have (+) or (–) values. Rather, it is the mere *presence* of a feature in the representation of a segment that designates the active involvement of that articulator (or acoustic feature) in the production of the segment; likewise, the *absence* of a feature entails that the articulator is not active for a given segment. There are strong theoretical arguments to support the hypothesis that all features are monovalent. From the viewpoint of parsimony alone, a system in which *all* features have the same number of possible values is to be preferred – whether that is one or two values. It has been shown convincingly, however, that certain features have only one value. Hence, keeping parsimony considerations in mind, the null hypothesis is that *all* features have only one value.

1.2.2. Underspecification

A closely related aspect of one's theory is the degree to which segments are specified (or underspecified). Three versions of underspecification have been advanced in recent years (e.g. Archangeli, 1984, 1988; Avery & Rice, 1989; Pulleyblank, 1986; Steriade, 1987). In this paper I adopt Avery & Rice's theory of underspecification, which can be labeled Minimally Contrastive Underspecification .

Minimally Contrastive Underspecification is inventory driven; that is, which features a segment is specified for depends *entirely* on what it contrasts with in the particular inventory. Segments are only specified for those features which are necessary to distinguish them from other phonemes in the language. According to Avery & Rice, a theory of universal markedness (UMT) is part of Universal Grammar. This UMT provides information as to which features are (universally) present and/or absent in underlying representations. Many linguists believe that the coronal place of articulation is the universally unmarked place of articulation (e.g., Kean, 1975; Paradis & Prunet, 1989; papers in Paradis & Prunet, 1991). In terms of underspecification theory, this means that coronal segments do not contain the Coronal node in their underlying representation, rather they receive it through the application of default rules.[12] There is, however, one circumstance in which coronal segments may be specified with the Coronal, namely when a language makes a phonemic contrast within the class of coronal segments. For example, in English, the segments /s/ and /š/ are both coronal, yet they differ with respect to their coronality; in particular, /š/ is further specified for the feature [+distributed].[13] Under Minimally Contrastive Underspecification, "if a secondary content node is the sole distinguishing feature between two segments (here [distributed]), then the primary feature (here Coronal) must be present in underlying representation" (Avery & Rice 1989:183). Thus, the representations of both /s/ and /š/ will contain a Coronal node, but only the /š/ must be further specified with the feature [distributed], as shown below.

(14)	/s/	/š/
	ROOT	ROOT
	|	|
	Place	Place
	|	|
	Coronal	Coronal
		|
		distributed

With these assumptions regarding feature specification outlined, I turn now to a discussion of the phonological operations that I will utilize.

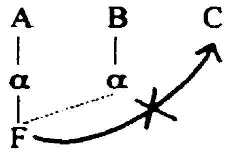
1.2.3. Phonological Operations

Given the current shift in phonological theory away from rules toward representations, the conception and formulation of phonological operations (e.g. Spreading, Fusion) are

[12] I will not address the issue of when these default rules apply. For arguments that these default rules apply only at the level of phonetic implementation, see Rice & Avery (1989, 1991b).

[13] It is not crucial here which precise coronal-dependent feature is posited to distinguish /s/ and /š/.

becoming an increasingly integral component of one's theory. The theory of Spreading that I will adopt can be thought of as a "feature filling" operation, similar to versions espoused by Piggott (1992, 1993) and Rice & Avery (1991b). It is schematized in (15), where A, B and C are organizing nodes of the same type (e.g. Place nodes), α represents content nodes of the same type (e.g. Labial) and F is some secondary feature (e.g. [round]).[14]

(15) 

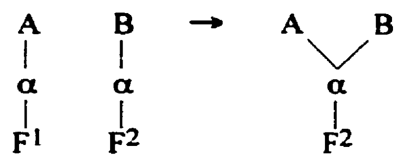
According to this formulation of Spreading, a feature may only spread to a segment not specified for that feature; encountering a segment that is already specified for that particular feature will arrest the operation. Importantly, Spreading does not trigger delinking. In addition, a feature may only spread to a segment if that segment contains a suitable docking site for the feature. In other words, the target segment must contain the superordinate node of which the spreading feature is a dependent. This theory of Spreading crucially does not allow any type of node generation (cf. Archangeli, 1984, 1988; Pulleyblank, 1986). Segments not containing the appropriate superordinate node will not be "visible" on the given tier. Those segments will, therefore, be transparent to the Spreading operation; they will neither be affected by the operation nor arrest the operation. Thus, cases of assimilation provide good evidence for the representation of segments; by the behavior of a segment in these processes, we can infer the presence or absence of particular features in that segment's internal structure.

Given the theory of monovalent features outlined above, it is clear why the operation of Spreading can only be "feature-filling." If there are no longer (+) and (–) values of a given feature on a tier, there will never be two different features (i.e. values

[14] Primary content nodes can, of course, also spread. Whether or not organizing nodes can spread is not as clear. In any case, according to the theory of Spreading that I am elaborating here, an organizing node such as the Place node could only spread to a segment not specified for a Place node at all (e.g. glottal stop).

of features) on the same tier. Assuming that a feature can only spread on its particular tier, then that feature will never encounter a feature that it could potentially change. Yet sometimes it appears that a segment may gain a different secondary feature in the process of assimilating to an adjacent segment that contains that particular feature. For example, we will see in the coronal harmony data below, that an anterior fricative may lose its anteriority and gain the feature [distributed] to match the place of articulation of a distributed sound in the string. Thus, it is clear that we need an operation that "changes" features. This "feature changing" spreading now falls under the rubric of Fusion.[15] This operation reduces a sequence of adjacent elements to just one. It may, in fact, effect a feature change of one of the elements. This operation is represented in (16); F^1 and F^2 represent different features (e.g., [anterior] and [distributed]).

(16)



Fusion is driven by the Obligatory Contour Principle (OCP), a well-established phonological principle. The separation of "feature-filling" and "feature-changing" operations is one of the ways in which phonological theory has shifted its focus from stipulative rules to phonological representations. The environment in which each of these operations can be executed is far more restrictive, making them more principled.[16]

---

[15] Not all linguists agree that Spreading can only "fill" a feature. Some linguists, in particular those that maintain that features are bivalent (e.g. Steriade, 1987), still consider a "feature-changing" operation to be Spreading, rather than Fusion. One of the reasons for rejecting the "feature-changing" type of Spreading is that it is far too powerful; if spreading can occur between two segments that are already specified for features on a given tier, then we allow for the possibility that this type of spreading will occur quite wildly. This type of spreading does not, however, occur in languages. "Feature-filling" spreading, in conjunction with the theory of underspecification assumed here, makes specific predictions as to which types of phonological processes will and will not occur. Moreover, this theory ensures that phonological processes occur as the result of representations, rather than stipulated rules.

[16] At first glance, Fusion appears to be as (if not more) powerful as the type of "feature-changing" Spreading that I just rejected. Fusion must, therefore, be constrained somehow. Here, I would like to draw upon a distinction between organizing nodes and content nodes proposed by Rice & Avery (1991b). They suggest that these two types of nodes are different in nature. One of the ways in which they differ is that organizing nodes are inherent to each segment and are specified individually for each segment (i.e. they may not be shared by segments); content nodes, on the other hand, may be shared, as the result of Spreading or Fusion. I would like to extend this distinction between organizing

Superficially, however, Fusion and Spreading look quite similar. Both are "directional"; whereas Spreading is leftward or rightward, Fusion may be left-headed or right-headed. The headedness of the Fusion operation determines which dependents of the node being fused will be retained. If the operation is left-headed, the dependents of the left most segment undergoing Fusion will be retained, supplanting the dependent features of the other segment(s) being fused. In addition, Fusion, like Spreading, requires locality: the two nodes being fused must be adjacent on the relevant tier. Thus, segmental processes that result from Fusion also provide good evidence for the representation of segments. With this theoretical background laid out, we can move on to examine the behavior of lateral segments in several different phonological processes.

1.3 Evidence For Unified Model

One type of segmental phenomenon that has been taken as evidence for the position of features in the geometry is harmony systems. Current autosegmental phonology analyzes harmony systems as cases of Spreading or Fusion, on a given tier. Therefore, harmony systems are actually cases of assimilation; the segments being affected are adjacent on the specified tier. In this section, I examine data from coronal harmony systems and nasal harmony systems, paying particular attention to the behavior of lateral segments.

1.3.1. Coronal Harmony

The first case I will outline is Tahltan Coronal Harmony. Tahltan is an Athapaskan language spoken in British Columbia. In this harmony, sibilants agree in anteriority

---

nodes and content nodes and suggest that organizing nodes do not participate in phonological operations while content nodes do. Drawing this distinction between these two types of nodes is a necessary step in order to constrain the Fusion operation. In particular, without this restriction to organizing nodes, the theory of Fusion outlined above allows for the Fusion of adjacent Place nodes (driven by the OCP), yet across-the-board Fusion of Place nodes whereby all of the segments of a string have the same place of articulation simply does not occur. An adequate theory will generate the attested possibilities as well as explain why other processes do not occur in languages. Restricting phonological operations to content nodes obtains the correct results. Rice & Avery assume a similar distinction; although, while they restrict operations such as Spreading and Fusion to content nodes, they incorporate an operation known as Copying which applies only to organizing nodes. The more restrictive theory constrains the application of *all* phonological operations to content nodes.

with the rightmost sibilant in the word. The data presented in this section come from Shaw (1991) who first used the Tahltan coronal harmony to argue against [lateral] as a dependent of the Coronal node. Tahltan has a rich inventory of consonantal phonemes, including five contrastive series of coronals, as shown in (17).

(17)	b	d	dl	dð̄	dz	dž	g	g^{w}	G	
		t	tl	tθ	ts	tš	k	k^{w}	q	
		t'	tl'	tθ'	ts'	tš'	k'	k^{w}'	q'	?
			ɬ	θ	s	š	x	x^{w}	χ	h
			l	ð̄	z	ž	ɣ	$ɣ^{w}$		
	m	n				y		w		
		n'								

The harmony system involves the three coronal series demarcated by the vertical lines above, (i.e. dð̄, dz and dž). It is worth mentioning that the lateral segments in Tahltan are fricative, not sonorant. This will become significant in the discussion below. Note also that this inventory contains lateral affricates in addition to plain coronal affricates.

Tahltan's coronal harmony has the following basic characteristics: 1) the harmony is directional, spreading from right to left, 2) the triggers and targets of the process include members of the dð̄, dz and dž series and 3) only the place of articulation, not the manner, spreads. This system is illustrated in the following examples. As shown in (18a–c), the first person singular subject marker, /s/, surfaces as [θ] if followed by any member of the dð̄ series, as [š] if followed by any member of the dž series, and as [s] elsewhere.

(18)	a.	/θεs̱ð̄εɬ/	→	[θεθ̱ð̄εɬ]	'I'm hot'
		/εs̱du:θ/	→	[εθ̱du:θ]	'I whipped him'
		/nas̱tθ'εt/	→	[naθ̱tθ'εt]	'I fell off (horse)'
	b.	/hudis̱tša/	→	[hudiš̱tša]	'I love them'
		/εs̱džlni/	→	[εš̱džlni]	'I'm singing'
		/ɬεnεs̱tšu:š/	→	[ɬεnεš̱tšu:š]	'I'm folding it'
	c.	/εs̱dan/	→	[εs̱dan]	'I'm drinking'
		/nadεdε:s̱ba:tɬ/	→	[nadεdε:s̱ba:tɬ]	'I hung myself'
		/sεs̱xεɬ/	→	[sεs̱xεɬ]	'I'm going to kill it'

In (19), the first person dual subject prefix /θi(d)/ surfaces as [š̲] or [s] in the appropriate context.

(19)	a.	/θ̲i:tθædi/ →	[θ̲i:tθædi]	'we ate it'
		/dεθ̲igɪtɬ/ →	[dεθ̲igɪtɬ]	'we threw it'
		/naθ̲iba:tɬ/ →	[naθ̲iba:tɬ]	'we hung it'
	b.	/dεθ̲it'us/ →	[dεs̲it'us]	'we are walking'
		/niθ̲it'a:ts/ →	[nis̲it'a:ts]	'we got up'
		/dεθ̲idzεl/ →	[dεs̲idzεl]	'we shouted'
	c.	/iθ̲itšotɬ/ →	[iš̲itšotɬ]	'we blew it up'
		/tεεdεnεθ̲idžu:t/ →	[tεεdεnεš̲idžu:t]	'we chased it away'
		/uθ̲idžε/ →	[uš̲idžε]	'we are called'

The first form in example (19b) illustrates that when another coronal fricative is present /θ/ loses its anteriority, thereby becoming an /s/. The coronal affricate also triggers this harmony as shown by the last two forms in examples (19b) and (19c).

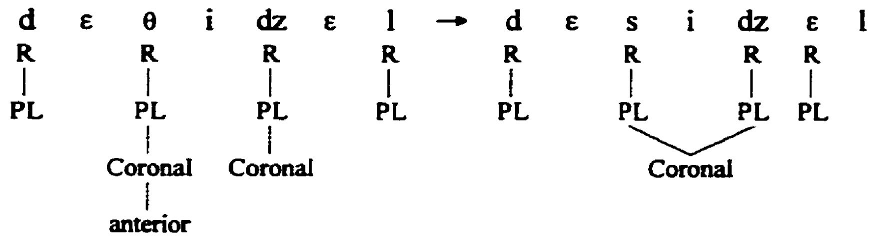
However, not all coronal segments trigger this harmony. For example, in the first form in (19a), the coronal /d/ does not participate – not because it is not a sibilant, but because it is not specified for a Coronal node. As the inventory in (17) indicates, the group of coronal stops, lateral fricatives and lateral affricates do not contrast with each other *within* the coronal place of articulation. Therefore, they need not be specified for any Place features. The fricatives /s/ and /z/, on the other hand, do contrast with the alveo-palatal fricatives. This is true of the affricates /ts/ and /dz/ as well. Thus, in accordance with Rice & Avery's Node Activation Condition (see §3.2), this contrast forces the presence of a Coronal node in each of these segments' representations.

This harmony, then, can be analyzed as right-headed Fusion of the Coronal nodes.[17] Fusing the Coronal node, rather than its dependents, is preferable since in some cases [anterior] is retained, sometimes [distributed] and, in some cases, a bare

---

[17] Operating on the Coronal node, itself, rather than its dependents accounts nicely for the fact that [anterior] or [distributed] may be transmitted to the relevant adjacent segments. However, one might argue that this Coronal harmony results from the *spreading* of the Coronal node, rather than the *fusion* of Coronal nodes. In certain cases (e.g. (18a)), Spreading would achieve the correct results. Nevertheless, I believe that there are very good reasons for maintaining that this Coronal harmony truly results from Fusion. In particular, spreading the Coronal Node violates the theory of Spreading that I outlined above. (e.g. it would be 'feature-changing' in that segments already specified for a Coronal node would be targeted). Also, it would require the additional operation of delinking to account for cases such as (19b) where an anterior fricative (θ) becomes a plain coronal (s).

Coronal node (i.e. no dependents) is retained. Fusing the Coronal node allows one to unify all of these processes under one operation. Example (20) shows a derivation of a form from (19b). The internal representation will be given only for the relevant segments; because manner features do not spread, only the Place node will be elaborated.

(20)


The two adjacent Coronal nodes are fused; because the operation is rightheaded, the features of the rightmost segment are retained – in this instance, a bare Coronal node.

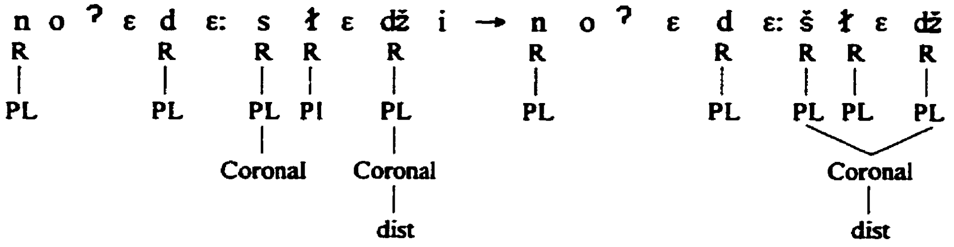
The striking aspect of this harmony is that both lateral fricatives and lateral affricates are completely transparent to the harmony: they do not trigger it, as shown by the forms in example (21a), and they neither undergo nor block the harmony, as shown by the examples in (21b).

(21) a.	/nεs̱tεɬ/	→	[nεs̱tεɬ]	'I'm sleepy'
	/sεs̱xεɬ/	→	[sεs̱xεɬ]	'I'm going to kill it'
	/nadεdε:s̱ba:tɬ/	→	[nadεdε:s̱ba:tɬ]	'I hung myself'
	/naθ̱iba:tɬ/	→	[naθ̱iba:tɬ]	'we hung it'
	/iθ̱itšotɬ/	→	[iš̱itšotɬ]	'we blew it up'
b.	/noˀεdε:s̱ɬεdži/	→	[noˀεdε:š̱ɬεdži]	'we melted it over and over'
	/yas̱tɬ'εtš/	→	[yaš̱tɬ'εtš]	'I splashed it'

The derivation below shows that the transparency of these lateral segments is captured nicely if we assume that they lack Place features.[18]

---

[18] If plain coronals, lateral fricatives and lateral affricates lack Place features as these Tahltan data suggest, then they must be distinguished from each other by some other means. As mentioned in footnote 7, with Steriade's (1993) proposal, fricatives can be distinguished from stops and affricates by virtue of their single articulatory phase (i.e. lack of closure; A_f (fricative) vs. A_oA_{max} (stop) or A_oA_f (affricate)). Coronal stops can also be distinguished from lateral affricates using these representations, by virtue of their respective *type* of release: stops have an "approximant" release, while affricates have a "fricative" release. Thus /t/ and /tɬ/ would have the following representations:

(22) 

If lateral segments are indeed unspecified under the Place node (and, therefore, transparent), we might expect other segments which are typically represented with a bare Place node, such as /t/ and /n/, to also be transparent to this Tahltan harmony. The forms in (23) demonstrate that this is, in fact, what we find.

(23)	/ɛdɛdɛs̱du:θ/	→	[ɛdɛdɛθ̱du:θ]	'I whipped myself'
	/tas̱tθaɬ/	→	[taθ̱tθaɬ]	'I'm dying'
	/xaʔɛs̱t'aθ/	→	[xaʔɛθ̱t'aθ]	'I'm cutting my hair off'
	/dɛθ̱it'təs/	→	[dɛs̱it'təs]	'we are walking'
	/niθ̱it'a:ts/	→	[nis̱it'a:ts]	'we got up'
	/mɛʔɛθ̱it'otš/	→	[mɛʔɛš̱it'otš]	'we are breast-feeding'

We can see here that /t/ and /n/ are not affected by the Fusion operation, and that /t/ does not block it. Given that lateral segments pattern in this coronal harmony with segments which are widely assumed to lack Place features, it is clear that the lateral segments also lack Place features.[19]

Before we set these Tahltan coronal harmony data aside, recall that the lateral segments in this language are non-sonorant fricative laterals. In this light, these data are

---

The /t/ has a coronal closure (underspecified) followed by an approximant (or maximally open) release whereas /tɬ/ has a coronal closure (underspecified) followed by a fricative release (also underspecified for Place) which is realized as a lateral fricative, as expected given my representation of a lateral fricative. However, a caveat is required here. One tenet of Steriade's proposal is that all segments contain one aperture node *underlyingly* and releases are projected (based on the observation that releases are not contrastive), thus /d/ and /dl/ should be indistinguishable underlyingly. Clearly, though, given my representations of laterals, stops must be distinguished from lateral affricates either in the way depicted above (with differing release types underlyingly) or, in traditional feature-geometric terms, as simple vs. complex segments (e.g. Hualde, 1988; Lombardi, 1990; Sagey, 1986).

[19] Labial and velar segments are also transparent. So, these data do not distinguish definitively between a model in which laterals have no Place features and a model in which [lateral] is a *direct* dependent of the Place Node (see §4.3.1 for arguments against the latter analysis).

even more informative, since they provide evidence to suggest that even non-sonorant laterals do not contain a Coronal node in their representation. Similar evidence comes from coronal harmonies found in Navajo, which also has a non-sonorant lateral and Chumash, which has sonorant lateral segments in its inventory. For the sake of space, I will not discuss the Navajo or Chumash facts here, but see Sapir & Hoijer (1967) for the Navajo data and Shaw (1991) for the Chumash data and analysis. The important thing to note here is that the sonorant laterals in Chumash and Navajo are transparent to the coronal harmony, just as the laterals in Tahltan. Thus, we can conclude that both sonorant and non-sonorant laterals are not *universally* specified for Place features. The behavior of laterals that we have seen in this section argues against any representation of laterality that requires a Coronal node (e.g. the Levin, Yip and Piggott models). Instead, these data provide support for the representations I am proposing. By examining languages with both types of lateral segments, we ensure a more accurate picture of the behavior of laterals and, hence, of their true representation.

1.3.2. Nasal Harmony

The coronal harmony we have just reviewed has allowed us to examine the Place features of lateral segments. We turn now to nasal harmonies, which will provide some insight into the manner features of lateral segments. Nasal harmonies crucially affect the sonorant features of a segment. Thus, the interaction of lateral segments with nasalization will indicate which sonorant features lateral segments are specified for; that is, what aspects of laterality are dependent upon specification for sonorant features.

Finding data to help sort out this debate is not easy; many of the languages that exhibit nasal harmonies lack lateral segments. Thus, the discussion here will be brief and the conclusions reached tentative. More concrete conclusions await additional data. I will present data from two languages which argue against the existing models of lateral representation and support the present analysis. While we saw no difference between the behavior of sonorant and non-sonorant lateral segments in coronal harmonies, suggesting similar representations under the Place node, with respect to nasal harmonies, we will see a difference between the behavior of these two types of lateral segments, suggesting dissimilar representations in terms of their manner of articulation.

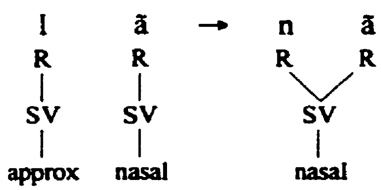
Itsekiri

The first nasal harmony that we will consider is exhibited by Itsekiri, a language spoken in the Mid-Western State of Nigeria. I will assume Piggott's (1992, 1993) analysis of this nasalization process as the result of Tautosyllabic Voice Fusion. This operation fuses adjacent SV nodes within a syllable. The SV features of the nucleus are dominant and, therefore, transmitted to the entire syllable, creating either an oral or a nasal syllable. In Itsekiri, approximant laterals are targets of nasalization and become full nasal stops as a result of this process. The data in (24) illustrate this.

(24)	lã	→	[nã] (*l̃ã)	'ask the price of'
	lɔ̃	→	[nɔ̃]	'be lost'

From these data we can see that the representation of laterals must be such that nasalization supplants the laterality of the segment, creating a full nasal, rather than a nasalized lateral. The derivation in (25) demonstrates the operation of Tautosyllabic Voice Fusion. The representation of the lateral here reflects my proposal.

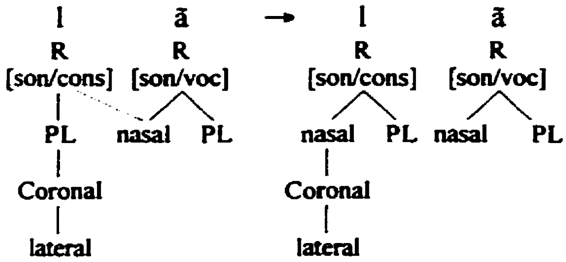
(25)



Tautosyllabic Voice Fusion fuses the SV nodes; since the nucleus is the head, the nasality of the vowel is retained, thereby creating a nasal syllable. The manner features of the lateral segment (i.e., [approximant]) are supplanted by the features of the head, which creates a full nasal. The laterality of this segment can only be supplanted by nasalization if we assume that laterality derives (in part, at least) from the specification for [approximant], which is a dependent of the SV node. Further evidence to support my proposal is provided by the nasal harmony in Yoruba. The full set of data are contained in Piggott (1993). As in Itsekiri, sonorant lateral segments in the Yoruba nasal harmony become full nasal stops as the result of nasalization.

These data show the Levin model to be inadequate for its inability to derive full nasal stops from lateral segments. Because Levin's model represents laterality distinct from the node involved in nasalization (i.e. SV), the resulting segment will always be specified for the features [nasal] *and* [lateral]. Therefore, this type of model predicts that sonorant lateral segments could only become nasalized laterals, not full nasal stops, as a result of nasalization.[20] The following derivation reflects her model.

(26)



Nasalized laterals, however, are not attested in Itsekiri or Yoruba. Since nasalization supplants the laterality of a segment, it appears that laterality (at least of sonorant lateral segments) must be represented as a dependent of the SV node.[21] In terms of the model advanced here, laterality in sonorant laterals must *derive* from specification for particular sonorant features.

While examining the behavior of sonorant lateral segments with respect to nasal harmonies is informative, we also need a study of the behavior of *non-sonorant* laterals

---

[20] The correct form can only be achieved by the addition of a questionable co-occurrence constraint that would delink the feature [lateral] once the feature [nasal] spread to the lateral segment. Such a constraint is unmotivated, however, since nasalized laterals are phonetically possible and found across languages.

[21] The language Gbe provides an apparent counter-example to my claims here. Laterals in this language do, in fact, nasalize (rather than become full nasals) as a result of the nasal harmony. Piggott (1993) analyzes this nasal harmony as the result of Tautosyllabic Voice Fusion and obtains the facts by positing that the lateral segments in Gbe, in contrast to those in Yoruba and Isekiri, are not consonantal. According to Piggott (1993), only consonantal segments may not be specified for both [approximant] and [nasal]. By virtue of being [-consonantal], the Gbe lateral segments are permitted to retain their feature [approximant], thereby becoming a nasalized lateral segment rather than a full nasal stop. The reader is referred to Piggott (1993) for a complete analysis of nasalized laterals and nasalized semi-vowels.

in nasal harmonies. A strong prediction can be made regarding these non-sonorant laterals: lacking an SV node, these segments should be entirely transparent to SV processes, such as nasalization. In the next section, we will see that this is indeed the case.

Slave

Slave is an Athapaskan language spoken in the Northwest Territories of Canada. This language, which is made up of four dialects (Slavey, Mountain, Bearlake and Hare), has been analyzed in great detail by Rice (1989). The nasal harmony data presented below come from this source. For our purposes, it is significant that the lateral segments in Slave are not sonorant, they are fricative.[22] Thus, this language presents a unique opportunity to investigate the interaction between non-sonorant lateral segments and nasal harmony.

Before we can examine the behavior of lateral segments, however, we must first review the Slave nasal harmony. This system has the following basic characteristics: 1) vowels are nasalized when followed by a nasal segment in the coda position of that syllable and 2) nasal segments in the onset position of the syllable are de-nasalized (resulting in a pre-nasalized stop) if no nasal segment follows in that syllable.[23] These characteristics are exemplified in the following forms (from Rice 1989:58-59, 144, 967). Tone will be omitted from all of the Slave examples as it plays no role in the nasal harmony.

(27)	a.	/-kon/	→	[kõ][24 25]	'fire'
		/-min/	→	[mĩ]	'net'

---

[22] This classification of the lateral segment is based, in part, on its participation in voicing alternations, paralleling fricatives in the language.

[23] Variation is found between a nasal and prenasalized stop; so for example, the forms /nétĩ/ --> [ndétĩ] are both acceptable. This suggests that nasal harmony in Slave (or in certain dialects) is optional.

[24] The obstruent /k/ in the (a) example is not affected by the Fusion operation because it lacks an SV Node.

[25] The nasal in the coda position of the (a) example is later deleted, producing [kõ]. We are assured of the presence of the nasal in the underlying form by forms with a vowel suffix, in which the nasal, no longer in coda position, is retained; for example, [kone].

b. /netĩ/ → [ndetĩ] 's/he went to sleep'
/themeh/ → [thembeh] 'it boiled'

This nasal harmony can be analyzed as the result of Tautosyllabic Voice Fusion, paralleling Itsekiri.[26] Positing the operation of Tautosyllabic Voice Fusion in Slave allows us to capture the nasalization of vowels and de-nasalization of nasals in a unified account. Thus, the operation of Tautosyllabic Voice Fusion is straightforward, creating oral or nasal syllables. With this analysis of Slave nasal harmony outlined, we can now consider the behavior of lateral segments in Slave.

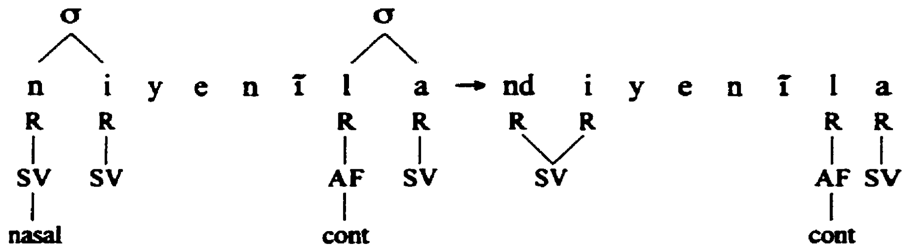
Recall that nasals are de-nasalized when they are followed by an oral vowel. The following forms reveal that lateral segments, both fricative and affricate (which appear underlined), are not in any way affected by an adjacent oral vowel (from Rice 1989: 58, 59, 63, 218, 968).

(28)	a. níyénĩ<u>ɬ</u>a	→	ndiyenĩ<u>ɬ</u>a	's/he placed them'
	b. -<u>tɬ'</u>u<u>l</u>é	→	<u>tɬ'</u>u<u>l</u>é	'rope'
	c. mĩ<u>l</u>e	→	mĩ<u>l</u>e	'net'
	d. nadéh<u>tɬ</u>ah	→	ndadeh<u>tɬ</u>ah	's/he started back'
	e. nõ<u>tɬ</u>ah	→	nõ<u>tɬ</u>ah	's/he made a return trip'

We must assume, of course, that Tautosyllabic Voice Fusion is also operative in the derivation of these words, even though we may not see any visible effects (e.g. nasalization or de-nasalization). Thus, the SV nodes of each syllable will be fused. The derivation in (29) illustrates this operation; the structure of the initial nasal is given for comparison with the lateral

---

[26] Keren Rice (p.c.) has brought a potential problem with this analysis of Slave nasal harmony to my attention. The issue Rice raises rests on the assumption that voicing is represented by the SV node; that is, voiced segments will have an SV node in their representations. I will not elaborate this view here, but see Avery (1993) and Rice (1993) for details. If we assume this to be the correct representation of voicing, then the following prediction is made: if nasal harmony results from Tautosyllabic Voice Fusion (which fuses SV nodes) then voiced segments should participate in this harmony. As Rice notes, voiced segments do not participate in nasal harmony in Slave. However, a few caveats are necessary here. First of all, voicing is not contrastive in Slave, so phonetically voiced segments will not be specified with an SV node underlyingly; thus, we would not expect them to participate in the harmony. Second, for those exceptional lexical items, which it appears voicing must be specified underlyingly, it is plausible that specification for continuancy blocks the participation of voiced continuants in the nasal harmony. Moreover, it is possible that the representation of voicing as an SV node is incorrect, in which case, the potential problem disappears.

(29) 

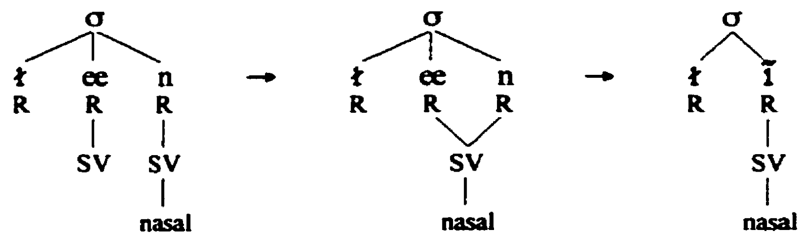
Note the difference between the behavior of the nasal in this example which becomes a pre-nasalized segment when followed by an oral vowel and the lateral which remains unaffected.

There is also evidence that the presence of an adjacent nasal vowel does not affect lateral segments; they become neither nasal stops nor nasalized laterals, as demonstrated by the forms in (30).

(30)

a.	hehłĩ	→	hehłĩ	'I am'
b.	lĩya	→	lĩya	'puppy'
c.	lõ	→	lõ	'end'
d.	gotl'ãa	→	gotl'ãa	'after'
e.	dlõa	→	dlõa	'mouse'
f.	/-łee + ñ/	→	[łĩ]	'be'

Example (30f) provides an excellent illustration of the fact that lateral segments in Slave are transparent to Tautosyllabic Voice Fusion; the vowel becomes nasal but the lateral segment is not altered. The representation given in (30) obtains the observed facts.[27]

(31) 

Based on the fact that lateral segments behave like obstruents (recall the form in (27a) /kon/ → [kõ]), we can be quite certain that fricative lateral segments, like obstruents, lack an SV node. Both the Rice & Avery and the Piggott model posit some feature as an

[27] The change of /ee/ to /i/ is an unrelated phenomenon discussed by Rice (1989) as Raising.

SV dependent to represent laterality in sonorant laterals. In order to capture the transparency of fricative lateral segments to this nasal harmony, laterals must be represented without an SV node. The representation of lateral fricatives proposed in this paper will account for the Slave facts.[28]

The behavior of non-sonorant laterals in nasal harmonies is especially interesting in light of the behavior of sonorant laterals we saw previously. It is clear that the laterality of non-sonorant laterals cannot be dependent on sonorant features or be derived in any way from sonorancy. Yet, many of the current models consider laterality of sonorant laterals to be inextricably linked to sonorancy (e.g. Piggott, Rice & Avery, Yip). However, if we want to maintain laterality as a uniform property, with similar representation in all types of lateral segments, then we must reject the notion that laterality is uniquely defined in terms of sonorant features.

1.3.3. Additional Evidence

We have seen that lateral approximants must be specified for sonorant features (in particular, [approximant] under the SV node) and that lateral fricatives and affricates cannot be specified for sonorant features. It still remains to be seen which manner features are specified for these non-sonorant laterals.

Navajo

I turn now to data from Navajo which suggest that fricative laterals have the same manner features as other fricatives. These data also give us some insight into the Place features of these segments. Navajo, as described by Kari (1976), has a rule of vowel deletion that creates a syllabic alveolar nasal. The following data from Navajo illustrate that there are some restrictions to this rule – it is obligatory before coronal stops,

---

[28] A model, like Levin's, in which the feature [lateral] is a dependent of the Coronal node will also account for the behavior of lateral segments in the Slave nasal harmony. However, as we saw in §4.1, Levin's model is unable to account for the behavior of non-sonorant laterals in other phonological processes, such as coronal harmonies.

optional before coronal fricatives and ungrammatical before non-coronals (Kari, 1976:92; glosses not given).

(32) a. coronal stops (obligatory)
nijigeeh → n̩jigeeh
nidoogoh → n̩doogoh

b. coronal fricatives (optional)
nisiłhiz → n̩siłhiz
niłchxo' → n̩łchxo'

c. non-coronals (ungrammatical or questionable)
nimá → *n̩ma
nahałtin → *n̩hałtin
niʔeeš → *n̩ʔeeš
nikʔaaš → ?n̩kʔaaš

The interesting aspect of these data is that the lateral fricatives pattern with the other coronal continuants. That this vowel deletion rule is optional before lateral fricatives indicates that they are specified for the same manner features as the other fricatives, namely [continuant]. These data also show that the lateral fricatives pattern with the class of coronals – the vowel deletion is grammatical before both types of segments.

These vowel-deletion data are compatible with a representation of non-sonorant laterals in which [lateral] is a dependent of the Coronal node (a reasonable extension of Levin's model); this type of structural dependency would explain why lateral fricatives pattern with the other coronal fricatives. However, since lateral fricatives are transparent to the coronal harmony in Navajo (Sapir & Hoijer, 1967), as they are in Tahltan (see §4.1), the hypothesis that [lateral] is a dependent of the Coronal node cannot be maintained.

Recently, Rice (p.c.) has suggested that laterality of non-sonorant laterals might be represented by some feature (perhaps [lateral]) which was a *direct* dependent of the Place (not the Coronal) node. If laterality of lateral fricatives were represented as an additional dependent of the Place node, then we would expect these segments to pattern with the non-coronals and vowel deletion before them should be ungrammatical. But as the data attest, deletion is grammatical before lateral fricatives. If laterality of non-sonorant segments is represented by some dependent of the Place node, there is no *a* *priori* reason for these segments to behave as coronals. Thus, any analysis that posits a feature [lateral] (or any other feature for that matter) as a dependent of the Place node must have some means of limiting this additional place feature to the coronal place of articulation. This co-occurrence restriction casts doubt on the plausibility that this "lateral" feature is a dependent of the Place node. Thus, these Navajo data are important because they give us insight into both the Manner and the Place features of non-sonorant laterals. The behavior of lateral fricatives as coronals can only be captured if they are represented with a coronal place of articulation or, in line with underspecification theory, a bare Place node. The representation that I am proposing for non-sonorant laterals would predict precisely the type of behavior that we see here in Navajo.

Throughout this paper I have emphasized that a definitive representation of lateral segments will account for the behavior of both sonorant and non-sonorant lateral segments. In the sections on coronal harmony and nasal harmony we examined data from languages with sonorant lateral segments and languages with non-sonorant lateral segments. It would also be informative to investigate languages that have both types of lateral segments in their phonemic inventories. Examining the behavior of both types of lateral segments in the *same* language will allow us to make an even more direct comparison of the two types of laterals.

Nez Perce

Nez Perce provides another opportunity to examine the place features of lateral segments. These data are especially important since Nez Perce contains both the lateral approximant and the lateral fricative in its phonemic inventory.[29] We can be certain that

[29] One might argue that what Aoki and Maddieson call a voiced lateral approximant is actually just the voiced counterpart of the voiceless lateral fricative; a voiced lateral fricative and a voiced lateral approximant are quite similar phonetically and many languages have both a voiced and a voiceless lateral fricative, with no lateral approximant. However, voicing is not contrastive in Nez Perce, so it is unlikely the approximant is actually a voiced counterpart of the voiceless fricative. Furthermore, the two segments behave quite differently phonologically. The lateral approximant patterns with the other sonorants in the language; for example, all nasals and semivowels (but not fricatives) become voiceless word-finally. The lateral approximant also becomes voiceless in this position. The voiceless fricative, on the other hand, alternates with a homorganic affricate between vowels. Thus, there is considerable

these two laterals differ in their manner of articulation (rather than just their voicing) based on their respective phonological behavior: the lateral approximant patterns with the other sonorants in the language in de-voicing word-finally; the voiceless fricative, on the other hand, alternates with a homorganic affricate between vowels. Moreover, voicing is not contrastive in Nez Perce, so it is unlikely the approximant is actually a voiced counterpart of the voiceless fricative. In our discussion of Nez Perce, we will focus on an assimilation process within the place of articulation. According to Aoki (1970), the dental consonants of Nez Perce, namely /t/, /n/ /l/ and /ɬ/ are palatalized before the vowel /u/. The forms in the following example illustrate this process (from Aoki, 1970:11).

(33)	a	t̪ʸu:skex	'upward'
	b.	peˀ t̪ʸú:qes	'man-crazy'
	c.	n̪ʸú:sn̪ʸu	'nose'
	d.	ˀil̪ʸú:t	'belly'
	e.	ɬ̪ʸúk'	'thump'

I will not attempt to characterize this process formally as it is not clear what the precise operation is. Notably, this "palatalization" effect is not found before the vowel /i/; it may, therefore, actually be velarization. Nevertheless, regardless of the precise formulization of this process, the fact remains that the sonorant lateral and the non-sonorant lateral are affected identically by this process. This suggests that their representation under the Place node is identical, as I am proposing.[30]

Welsh

I would now like to consider a bit of data from the acquisition of lateral fricatives and lateral approximants. Welsh has both of these types of lateral segments, so it should prove to be informative. The acquisition data that we will be examining are reported by

---

motivation for maintaining that these two lateral segments have different manners of articulation (i.e., that one is an approximant and the other is a fricative).

[30] Although these data are compatible with Levin's model (if we extend her analysis to non-sonorant laterals), we have seen several cases which show clearly that laterals (both sonorant and non-sonorant) cannot contain [lateral] as a dependent of the Coronal node (e.g. Chumash Korean, Navajo, Tahltan).

Bellin (1984). The author notes that Welsh children have problems acquiring the lateral fricative and attributes this difficulty to a more general difficulty in acquiring fricatives cross-linguistically. It is quite interesting that the difficulty in acquiring these lateral fricatives does not stem from a difficulty with the articulation per se; the children were able to produce the sound, but only as a deformation of a different adult form (such as tw).

In acquisition research, the types of phonological substitutions a child makes are often quite indicative of her internal representation of the replaced segment (i.e. the one that is substituted for). For example, it is argued by some that coronal segments are the segments most often substituted because they are have an unmarked place of articulation (Stemberger & Stoel-Gammon, 1991). The substitutions made by the Welsh children for lateral fricatives and for lateral approximants should tell us something about the basic character of each of these segments. Bellin reports that the most common substitutions for the lateral fricative were /fl/, /x/, /sl/ and /θ/. These substitutions indicate that the child has determined one of the primary properties of lateral fricatives, namely continuancy; all of the substitutions are, themselves, continuant. This would suggest that the feature [continuant] is part of the representation of non-sonorant laterals. It appears, though that the child has not mastered the place of articulation; each of the substitutions reflects a different place of articulation.

If sonorant laterals are defined by some combination of features that is different than those present in a fricative lateral, then we should expect the Welsh children to use different substitutions for the two types of laterals. Unfortunately, Bellin does not report any cases where the sonorant lateral is substituted for. However, it is reported that the sonorant lateral is, itself, often substituted for the approximant segment /w/. It appears that the children have not mastered labial approximants; their substitution is a coronal segment. Assuming that like segments are normally substituted for another segment (e.g. fricatives were substituted for fricative laterals above), we can infer that the representation of sonorant laterals in Welsh contains specification for [approximant]. Again, it appears that place of articulation is not a factor in these Welsh substitutions.

What conclusions, then, can be drawn from these acquisition data? I think these data demonstrate the basic nature of lateral segments. According to Rice & Avery (1991a), children, in acquiring their language's phoneme inventory, first detect contrastive use of segments and then elaborate (in their mental grammars) the structure differentiating these phonemes. The make-up of a segment consists primarily of two things - manner features and place features. Children detect contrasts along these two spectrums; that is, they will differentiate segments either in terms of manner features or place features, initially. The data reported above indicate that the Welsh children are, at this stage in their development, contrasting lateral segments with other segments in terms of manner, not place, features. Thus, we may take the substitutions to reveal the basic nature (i.e. features) of the target segment – [continuant] in the case of non-sonorant laterals and [approximant] in the case of sonorant laterals.

1.4 Review of Existing Data in the Literature

A large amount of data has already been brought to bear on the issue of lateral representation. Levin, Rice & Avery and Yip have amassed a considerable amount of data in support of their respective models. A critical test, then, for the model of lateral representation I am advocating is whether or not it can account for the data previously brought to bear on this issue, in addition to the data I have presented here.

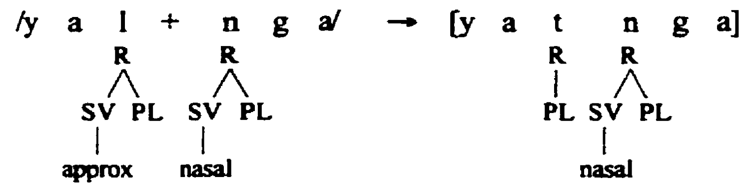
Yip (1990) discusses ten primary cases that have been used to determine the representation of lateral segments. All of these languages involve sonorant, rather than non-sonorant, lateral segments. I will briefly consider each of these languages and show how my proposal accounts for each one. The first set of languages I will comment on have been used by Rice & Avery to support a model in which [lateral] is a dependent of the SV node. Because my representation of sonorant laterals differs very little from theirs (i.e. both of us assume an SV node and an SV dependent, as well as no Place features) little discussion will be required. For the sake of spacc I will only present enough data to illustrate the basic phenomenon. For further examples and discussion the reader is referred to the relevant sources.

Kuman is one of the languages that Rice & Avery (1991b:111) cite to support their model. The following data illustrate a process of desonorantization in which a lateral followed by a nasal becomes a coronal stop.

(34) a.	/yobuḷ + na/	→	[yobuṭna]	'my bone'
b.	/yaḷ + nga/	→	[yaṭnga]	'you plant'

Rice & Avery analyze this process as delinking of the SV node, which is driven by the presence of another adjacent SV node. The derivation of (34b) is given in (35), using the representation of laterality advocated in this paper.

(35)

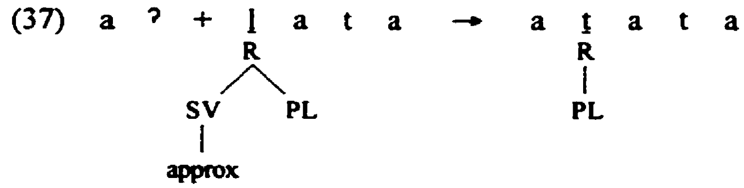


The proposed representation easily accounts for these data because, like Rice & Avery, I assume that one of the features necessary in the representation of sonorant laterals is a dependent of the SV node. Once the SV node is delinked, one of the features necessary for the realization of laterality (namely [approximant]) is missing and the result is a coronal obstruent.

Data from the Move dialect of Yagaria show the same type of alternation between sonorants and obstruents; the sonorant form occurs after a vowel and the obstruent form occurs after a glottal stop (which is the only syllable-final consonant, and is subsequently lost). This alternation is shown below (from Rice & Avery, 1991b:110).

(36)	bade	+	ḷata	→	badeḷata	'two boys'
	aʔ	+	ḷata	→	aṭata	'two women'

The following derivation illustrates how this alternation can be captured without using the feature [lateral].

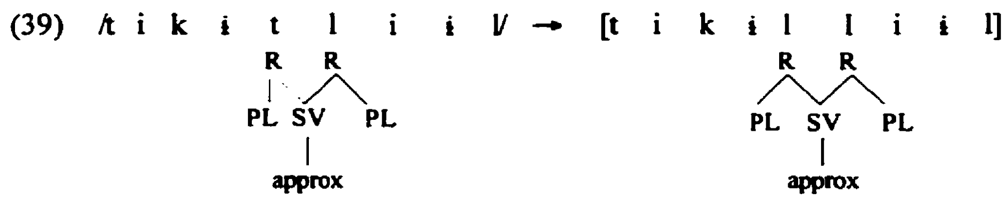


The SV node of the lateral is delinked (conditioned by the obstruent in the preceding coda), resulting in a coronal obstruent. It is also of interest that although laterals in Yagari pattern phonologically with the coronal segments (Levin, 1988), they are phonetically velar (Renck, 1967). This is unexpected if laterals are universally specified with a Coronal node. However, if we assume that laterals lack Place features (unless required by underspecification theory), then we allow for the possibility that some languages will realize the bare Place node with a velar, rather than coronal, place of articulation (see Trigo, 1988 and Rice, 1992 for more on velar place of articulation resulting from an unspecified Place node)

Rice & Avery have also marshaled the Korean data discussed above in §1.1 in support of their model. These data from (1), which illustrate a process of sonorantization, are repeated here in (38).

(38) a.	/kukmul/	→	[kuŋmul]	'soup'
b.	/kakmok/	→	[kaŋmok]	'wood'
c.	/napnita/	→	[namnita]	'to sprout'
d.	/kat^{h}ni/	→	[kanni]	'to be the same'
e.	/tikitlɨɨl/	→	[tikillɨɨl]	'the letters t and l'

This process, by which stops become nasals (a–d) or laterals (e), can be analyzed as leftward spreading of the Sonorant Voice node.

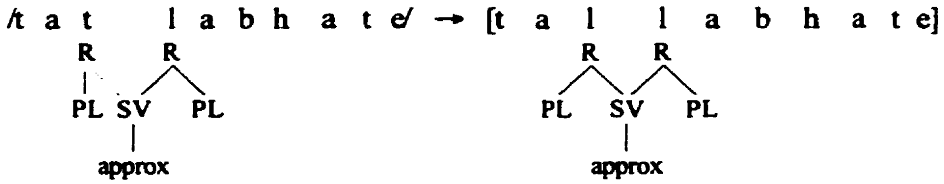


The derivation of (e), employing the proposed representation of laterality, is shown above in (39) (the identical Place nodes may be fused later).

The next case I must account for comes from Sanskrit, which has an assimilation process similar to Korean: stops assimilate to the nasality of a following segment and /t/'s can assimilate to a following lateral. The data are from Rice & Avery (1991b:113).

(40)	a.	/taṭ namas/	→	[taṇ namas]
	b.	/triṣṭup nunam/	→	[triṣṭum nunam]
	c.	/vak me/	→	[vaŋ me]
	d.	/taṭ labhate/	→	[taḷ labhate]
	e.	/uṭ luptam/	→	[uḷ luptam]

The process illustrated here, like the Korean data, can be analyzed as leftward spreading of the SV node; a derivation is provided in (41).

(41)


These data are also captured elegantly by the proposed representations: the original coronal stop now has an SV node and an [approximant] dependent and will, therefore, be realized phonetically as a lateral segment.

The next two cases that we will consider involve place assimilation. The logic behind these examples is that if [lateral] is indeed a dependent of the Place node then laterality should spread, just as other Place features do. Javanese is the first language that we will examine. The data are from Yip (1990:11).

(42)	a.	/ŋ + aḍjar/	→	[ŋaḍjar]	
	b.	/ŋ + bakar/	→	[mbakar]	'roast'
	c.	/ŋ + ḍuḍuʔ/	→	[ṇḍuḍuʔ]	'place'
	d.	/ŋ + lunguh/	→	[ŋlunguh]	'sir'

As examples (42a–c) show, the velar nasal assimilates in place of articulation to the following consonant, but fails to assimilate to a lateral segment, in example (42d). This fact can only be captured if we assume that laterals are not universally specified for

Place features; lacking Place features, there is nothing to spread from the lateral to the velar nasal. A derivation of (42d) is shown in (43); given the theory of Spreading as a "feature-filling" operation, we must assume that the velar nasal is unspecified for place features. Note that these data argue against the idea that lateral segments contain a Coronal node universally; if laterality required a Coronal node, we would incorrectly expect the velar nasal to assimilate to the coronal place of articulation, as it does before the coronal obstruents.[31]

(43) c. /ŋ + l u n g u h/ → [ŋ l u n g u h]

Dutch provides the second case of place assimilation. As the data in (44a–e) show, /n/ assimilates to the place of articulation of the following consonant, yet fails to assimilate to the laterality of a following lateral (44f–g) (Rice & Avery, 1991b; originally from Trommelen, 1984).

(44) a.	i[n] elf uur	'in eleven hours'
b.	i[m]–bringen	'to bring in'
c.	i[ŋ]–kopen	'to purchase'
d.	i[ɲ] jullie huis	'at your house'
e.	i[ɱ] vieren	'in four parts'
f.	i[n]–laten	'let in, admit'
g.	i[n]–leggen	'lay in, put in'

The failure of the nasal to assimilate to the laterality of /l/ can be explained by the representation of laterality advanced here: laterality derives from a combination of

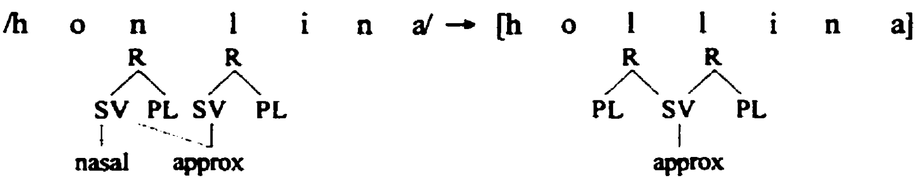
[31] Two points require comment here. First, coronal stops are specified for a Coronal node in Javanese since there is a contrast between dental and retroflex coronals, thus there is feature material under the Place node that is available for spreading. The second note is that /r/, like /l/, does not spread its place of articulation to the nasal, suggesting that it too is unspecified for Place features (e.g. ŋrantap-ake 'to flake off'). In §6.1, however, I will argue that the approximant /r/ is differentiated from the approximant /l/ by the presence of a coronal node in the representation of the /r/. We must keep in mind, though, that this representation holds only of the approximant /r/. The Javanese /r/ is, in fact, a trill and can be distinguished from the approximant /l/ by manner features (perhaps [continuant]). Thus, these data from Javanese do not disconfirm my proposal for the representation of the approximant /r/.

feature specifications, thus, spreading of laterality requires transmission of both sonorancy and coronality.[32]

The final two phenomena offered in support of Rice & Avery's model that I will examine come from Klamath and Ponapean. The phonological process at issue is identical in both of these languages, so I will treat them as a whole. The forms in (45) from Klamath and (46) from Ponapean exemplify the process, in which coronal nasals become lateral segments. As the (45c–d) examples show, non-coronal nasals and coronal non-sonorants are not affected (Rice & Avery, 1991b:108, originally from Barker, 1964:79 and Rehg & Sohl, 1981: 57).

(45)	a.	/honlina/	→	[hollina]	'flies along the bank'
	b.	/w'inl'ga/	→	[w'illga]	'lies down on the stomach'
	c.	/sl'oslq'a/	→	[sl'oslq'a]	'shed hair'
	d.	/t'omt'oml'i/	→	[t'omt'oml'i]	'flat on top'
(46)	a.	/nan-leng/	→	[nalleng]	'heaven'
	b.	/pan lingan/	→	[pallingan]	'will be beautiful'

In both of these languages, two adjacent SV nodes are fused and the result is a lateral segment. The following derivation shows how this is accomplished without the feature lateral.

(47)


It is clear why non-sonorant coronals and sonorant non-coronals (45c–d) are not affected: non-sonorants do not have an SV node, so the structural description of the operation is not met, and fusing the SV node of a non-coronal would produce an illicit labial approximant – only when a segment with both an approximant and coronal (or placeless) specification is created, as in the (a–b) examples, does a lateral result. This

[32] These Dutch data do not indicate whether or not lateral segments in this language are specified for a Coronal Node. The nasal preceding the lateral is coronal underlyingly and remains coronal, thus, there is no means of ascertaining if a Coronal Node or nothing at all was spread from the lateral.

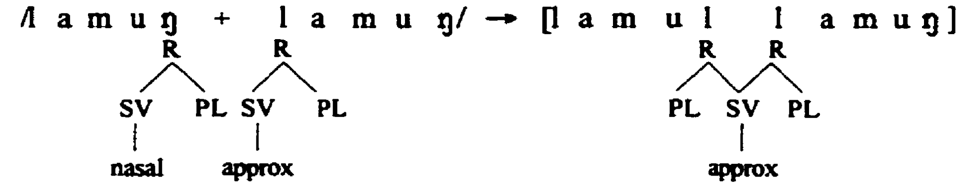
concludes my discussion of the data in the literature that have been used to support the Rice & Avery model. It is clear that my model can account for these data.

The remainder of the languages that will be examined in this section are all regarded as evidence for Levin's model of lateral representation. Levin (1988) uses primarily three coronal processes to argue for the dependency of the feature [lateral] on the Coronal node. Below, I review each of these arguments and re-analyze them in terms of the representations just proposed. Levin's strongest bit of evidence comes from Selayarese. In Selayarese, nasals assimilate to the point of articulation of any following consonant and completely to /l/. The process can be seen in the following reduplicated forms (Mithun & Basri, 1985).

(48)	a.	/dodoŋ + dodoŋ/	→	[dodondodoŋ]	'sort of sick'
	b.	/bambaŋ + bambaŋ/	→	[bambambambaŋ]	'sort of hot'
	c.	/gintaŋ + gintaŋ/	→	[gintaŋgintaŋ]	'chili-like object'
	d.	/soroŋ + soroŋ/	→	[soronsoroŋ]	'drawer'
	e.	/nunruŋ + nunruŋ/	→	[nunrunnunruŋ]	'hit lightly'
	f.	/maŋ ŋaŋ + maŋ ŋaŋ/	→	[maŋ ŋammaŋ ŋaŋ]	'sort of tired'
	g.	/ɲ amaŋ + ɲamaŋ/	→	[ɲ amaɲ ɲamaŋ]	'rather delicious'
	h.	/lamuŋ + lamuŋ/	→	[lamullamuŋ]	'plantation'

The significant aspect of these data is that what appears to involve only spreading of the Place node creates a lateral segment before another lateral segment (48h), suggesting that laterality is located under the Place node. While it is clear that there is a rule of place assimilation operating here, it is possible that sonorant features are also being transmitted from one segment to another (as demonstrated for Catalan in §1.1). Rice (1992) points out that languages do not allow consonants in the coda of a syllable to be less sonorous than the onset of the following consonant; coda–onset sequences that violate this sonority hierarchy are repaired by the application of various phonological operations (see Rice, 1992 on Ponapean, Toba Batak and Diola-Fogny). I would like to suggest that the spreading of laterality in Selayarese results from the spreading of sonorant features in order to create a permissible coda–onset sequence. Given that sonorant laterals are more sonorous than nasals, the derivation of (48h) would proceed as follows:

(49)



Right-headed fusion of the adjacent SV nodes ensures that the coda consonant is not less sonorous than the following onset consonant. The velar nasal is clearly underspecified in this language, so that acquisition of the feature [approximant] from the neighboring lateral creates a structure that will be realized as a lateral. Example (48h) above is the only case where we can see the effects of this operation; in examples (48a–d), the righthand member of the heterosyllabic cluster lacks an SV node, so that the structural description of the rule is not met (i.e. presence of two adjacent SV nodes), and in examples (48e–g), the two members of the heterosyllabic cluster are both nasal, so that the fusion of the SV nodes is vacuous.[33] Unfortunately, apart from total geminates, /ŋ/ and /ʔ/ are the only possible coda consonants in Selarayese, so we cannot determine the effect of the proposed SV fusion on an /l + n/ sequence or the effect of the place assimilation rule on a /t + l/ sequence. The introduction of an additional operation does make our account of the data in (48) a bit less elegant; however, if we assume that more is at play than only a rule of place assimilation, we are able to maintain the representation of laterality presented here.

One of Levin's more general arguments concerns languages in which /l/ alternates with coronals, sometimes acting as the [continuant] version of the coronal stops.[34] Yip (1990:10) cites Taiwanese as such a language, in which /t/ becomes [l] in a context where other stops simply voice. It is no accident that laterals alternate with coronal stops: the addition of the feature [approximant] or [continuant] to a segment

[33] A counter-example to the claim that laterals in Selarayese result from SV node Fusion rather than Place assimilation is /roŋgaŋ + roŋgaŋ/ → [roŋganroŋgaŋ] 'rather loose'. Here, the coda nasal assimilates to the coronal place of articulation of the /r/, but not to its sonorancy. At this point, I have no explanation for the failure of /r/ to transmits its sonorant features to the nasal.

[34] Yip (1990) does not indicate if this lateral is the sonorant or non-sonorant one. It would be especially interesting if this alternation occurred with the non-sonorant lateral, since it differs from coronal stops only by continuancy.

with a peripheral (i.e., labial or dorsal) place of articulation would not produce a phonetically lateral segment. The alternation between /t/ and [l] in a context where other stops simply voice (for instance, between vowels) can also be explained. If the sonorancy of a vowel is transmitted to a coronal stop, the result would be a lateral approximant. Thus, these sort of data provided by Levin in support of her model of lateral representation can actually be accounted for without the feature [lateral].

The last argument for Levin's model that I will consider here involves place assimilation in Basque. Basque has a process of place assimilation in which the nasal /n/ undergoes total place assimilation while /l/ remains coronal, although it takes on the appropriate values of [anterior] and [distributed]. The following forms illustrate this (Yip, 1990:10, originally from Hualde, 1988)

(50)	a.	egu[n] a	'the day'
		egu[m] berri	'new day'
		egu[m] fresken	'cool day'
		egu[n̪] denak	'every day'
		egu[ɲ] tiki	'small day'
		egu[ŋ] gorri	'red day'
	b.	ata[l] a	'the section'
		ata[l] berri	'new section'
		ata[l] fresken	'cool section'
		ata[l] denak	'every section'
		ata[ʎ] tiki	'small section'
		ata[l] gorri	'red section'

Levin argues that if [lateral] is a dependent of the Coronal node, then laterals must have Coronal nodes.

While these data do suggest that lateral segments in Basque do have a Coronal node, I do not think (as Levin does) that the presence of this Coronal node implies the presence of a [lateral] dependent. There is a perfectly good alternative explanation for the presence of a Coronal node in the representation of laterals. Many segments in Basque contrast within the coronal place of articulation, requiring that they be specified for a Coronal node.[35] Thus, the lateral segments are specified for a Coronal node as

---

[35] It is curious that /n/ undergoes total place assimilation while /l/ does not. Basque only contrasts labial, alveolar and palatal nasals (/ŋ/ is not contrastive). Thus, it is possible that /n/ and /ɲ/

required by underspecification theory, not because their laterality derives from, or depends on, the presence of this node.

The claim that only a Coronal node, rather than a Coronal node and a [lateral] dependent, is present in the Basque lateral is further supported by the following data (Yip, 1990:20).

(51)	egu[n] luze	'long day
	ba[t] luke	'he would have one'

If the feature [lateral] truly were a dependent of the Coronal node, then it should spread in this place assimilation process. Yet, as the forms in (51) show, it does not. The representation of lateral segments I proposed correctly accounts for the data: the Coronal node (with no dependents) of the lateral segments spreads leftward, resulting in a coronal segment. Thus, it is clear that the data from Basque do not support Levin's model and can, in fact, be accounted for by the model advanced here.

To conclude this section, I have shown that the model I am proposing is able to account for the data that has been amassed in the phonological literature around this lateral debate. I have by no means exhausted the data in the existing literature, only shown that the more prominent cases can be handled without the feature [lateral].

1.5 Implications

One of the basic claims of my proposal is that /l/ is the unmarked coronal approximant and /ɬ/ is the unmarked coronal fricative. By this, I mean that each is represented with the least amount of structure in their respective classes. While it is beyond the scope of this paper to fully investigate the representation of non-lateral segments, in the following section I take a brief look at some of the implications of my proposal for the representation of some closely related segments.

---

can be distinguished from each other without Coronal node dependents; in this case, /n/ would not need to be specified for Place features. This would explain why /n/ undergoes total place assimilation but /l/ does not.

1.5.1. The representation of non-lateral approximants

Non-lateral approximants, such as /r/ must somehow be distinguished from the lateral approximant. The representation that I propose for /r/ appears in (52).

(52) / r / (approximant)

ROOT

SPONTANEOUS VOICE PLACE

approximant Coronal

According to this structure, the only thing that distinguishes /r/ from /l/ is the presence of a Coronal node in the representation of /r/. It is very important to keep in mind that there are many varieties of r-sounds, ranging from trills to taps to approximants to semi-vowels. The structure in (51) is only intended to represent the approximant /r/, which would have the same manner features as the approximant /l/.[36]

There are two significant predictions that follow from the proposed representation in (52). One is that if both an approximant /l/ and an approximant /r/ are contained in an inventory, /r/ must always be specified for a Coronal node and will not be transparent to coronal processes such as coronal harmony. Unfortunately, none of the languages that I have examined with respect to coronal harmony in this paper contains an /r/ in its inventory. However, I know of two Australian languages, Gooniyandi and Diyari, which have coronal harmonies and contain /r/ in their inventories; in both languages, /r/ participates in the coronal harmony (Hamilton, 1993). Thus, there is preliminary evidence to support this prediction about /r/ in coronal harmonies. A second prediction is that in those languages where /r/ behaves as the unmarked (i.e. underspecified) approximant – for example, Japanese, as argued by Mester & Itô (1989) – there will not be a contrast within the class of coronal

---

[36] Clearly, an /r/ that patterns as a fricative or as a semi-vowel will be specified for the manner features that characterize fricatives or semi-vowels, respectively; in these cases, different specification for Place features would not be necessary to distinguish the trill or semi-vowel /r/ from the approximant /l/.

approximants (i.e. /l/ and /r/ will not contrast phonemically). We know that for Japanese, this is indeed the case.

1.5.2. The representation of non-lateral fricatives

Another implication of my proposal concerns the representation of the other coronal fricative /s/. This structure appears below.

(53)
```
          / s /
          ROOT
        /      \
   AIR FLOW    PLACE
       |          |
  continuant   Coronal
```

This coronal fricative is distinguished from the lateral fricative by the addition of a Coronal node, just as /r/ is distinguished from the lateral approximant. The claim that the lateral fricative, rather than /s/, is the unmarked coronal fricative may be controversial, since /s/ is regarded by many phonologists to be unmarked. Indeed, if /s/ is the *only* coronal fricative in a particular inventory, then it will be unmarked for Place features. Thus, the prediction of the above representation is that if an inventory contains both /ɬ/ and /s/, /s/ will be specified for a Coronal node, and will not be transparent to coronal processes. We saw that this was, in fact, the case in Tahltan, Navajo and Chumash. Many of the languages that have both /ɬ/ and /s/ also have /š/ in their inventories, so the presence of a Coronal node in the structure of /s/ is independently motivated.

This proposal also predicts that in languages that contain the lateral fricative, and no /s/, the lateral fricative should behave with the same extra freedom that /s/ enjoys in languages like English. Languages that contain /ɬ/ yet lack /s/ are exceedingly rare (a total of seven in Maddieson's (1984) *Patterns of Sounds*) – a fact that also deserves explanation – thus, whether this second prediction is borne out is still an empirical question at this point.

1.5.3. The acquisition of laterals and related phonemes

The final implication of my proposal that I will discuss concerns the acquisition of these segments. If /l/ and /r/ are distinguished solely by the presence of a Coronal node, then the prediction is that children will only contrast these two segments phonologically once they have posited the Coronal node in their grammar; that is, when they begin to contrast other sounds, such as /s/ and /š/, which also require a node to distinguish them. Recent work by Brown and Matthews (1993, in press[37]) on the acquisition of English phonemes provides support for the idea that /l/ and /r/ are distinguished by the presence of a Coronal node. The authors show that /l/ and /r/ are distinguished phonologically quite late (after sonorants are distinguished from obstruents and nasals are distinguished from non-nasals). Thus, it appears that whatever distinguishes these two segments is acquired relatively late. Interestingly, Brown & Matthews also found that coronal segments are the last to be distinguished from the other places of articulation, suggesting that the Coronal node is the last node under the Place node to be acquired. These acquisition data are compatible with the present proposal. Brown (1993a, to appear a, b[38], c) presents evidence from second language acquisition that also supports the proposed representations of /l/ and /r/.

Similarly, if /ł/ and /s/ are distinguished solely by the presence of a Coronal node, then children learning languages with these two phonemes (e.g. Welsh) should only contrast them once they have posited the Coronal node in their grammar. It remains to be seen if this prediction is substantiated. However, based on the production data from Welsh reported above, the prospects are promising, as children who do not distinguish segments in terms of their place of articulation do not contrast /ł/ and /s/ either.

---

[37] Brown & Matthews (in press) appears in this dissertation as Chapter 2.

[38] Brown (to appear a, b) appear in this dissertation as Chapters 4 and 3, respectively.

1.6 Conclusion

One of the goals of this paper has been to examine the phonological behavior of both sonorant and non-sonorant lateral segments and, thereby, obtain a more accurate picture of the *essence* of laterality. Following Spencer (1984) and, more recently, Piggott (1993), I have argued that [lateral] is best regarded as a phonetic, rather than phonological, feature. This claim, in conjunction with my proposed representations, provides an explanation for the propensity of lateral segments to have a coronal place of articulation – if a segment is specified for any Place feature (other than coronal, perhaps), it is simply not realized as a lateral. I have also argued that sonorant and non-sonorant laterals have similar representations. Until now, the internal structure of non-sonorant laterals has not been addressed. Importantly, the proposed representations for both types of laterals do not require the introduction of any new features, or dependency relations, into the overall feature geometry.

To support my proposed representations, I presented evidence from a wide range of segmental processes and a variety of languages. Using data from a coronal harmony data in Tahltan, as well as Navajo and Chumash, I demonstrated that laterals cannot be universally specified for Place features. This was taken to indicate that laterality cannot derive from specification for coronal features, thereby casting doubt on the models proposed by Levin, Yip and Piggott. Similar data from nasal harmonies was also examined. These data showed that while laterality must derive from specification for sonorant features for lateral approximants, it cannot for lateral fricatives. None of the models considered can account for the behavior of both types of laterals in these harmonies. The unified model presented here, on the other hand, is able to account for the behavior of both lateral types in the coronal and the nasal harmonies. To determine the structure of lateral fricatives, we reviewed additional data from Navajo, Nez Perce and Welsh. The behavior of lateral fricatives with respect to the rule of vowel deletion in Navajo suggested that they do contain a specification for the feature [continuant], as I proposed. Moreover, these data, as well as that from Nez Perce, demonstrated that the laterality of non-sonorant laterals cannot be adequately captured with an additional Place feature. The data from Nez Perce also confirmed that the two types of laterals have the same structure under the Place node. The acquisition data from Welsh gave us an indication of the basic nature of lateral approximants and lateral fricatives; in particular that sonorant laterals are specified for the feature [approximant] and that non-sonorant laterals are specified for the feature [continuant]. Thus, the data presented demonstrates that none of the existing models is able to account for the entire range of cross-linguistic data. At the same time, these data provide empirical support for the unified model advocated here.

The important and close relationship between the phonological component and the phonetic component of the grammar has, in recent years, received increased attention. The ways in which (and means by which) phonological structure is ultimately realized by speakers deserves systematic study. The claim that laterality is a phonetic property that results from a particular phonological structure underscores the close relationship between the two components. To truly understand the representation and realization of laterality we must now develop theories of its phonetic implementation.

Preface to Chapter 2

The system of mental representations proposed to characterize adult linguistic knowledge must do two things: 1) it must account for cross-linguistic data (i.e., what is grammatical and ungrammatical in specific languages, as well as in language in general), and 2) it must be learnable; that is, coupled with a theory of acquisition, a theory of grammar must explain how that linguistic knowledge is attained (Chomsky, 1965, 1981; Pinker, 1984). In the previous chapter, we considered the representation of laterality and examined how the feature-geometric representations posited for lateral fricatives and lateral approximants account for their behavior in several phonological processes cross-linguistically. Thus, the theory of Feature Geometry, and the feature-geometric representations proposed for lateral segments, are well supported by a range of cross-linguistic phonological data. However, if we assume that our theoretical characterizations of linguistic knowledge must be learnable by the child through exposure to the ambient language (i.e., there must be a way into the system for the child), then we must also determine whether (and how) the segmental representations posited within the theory of Feature Geometry, like those for lateral segments described in the previous chapter, can be acquired given the appropriate input. The theory of Feature Geometry is sufficiently developed at this point that we are now in a position to consider the predictions this theory makes for the acquisition of segmental representations.

The study of lateral segments in Chapter 1 has several important implications for the acquisition of these segments. First, we saw that theoretical considerations compelled the elimination of [lateral] from the set of universal phonological features. One of the consequences of removing [lateral] from phonological representations is that this feature is no longer presumed to be involved in the phonological systems of the world's languages, nor is it implicated in the development of any such system; in other words, [lateral] should play NO role in the acquisition of lateral segments or closely related segments, either in predicting the order in which these phonemes are acquired or in phonological rules posited by children. Eliminating [lateral] from the phonological system also requires that lateral segments be distinguished from other related segments (i.e., approximants or fricatives) in a speaker's mental grammar by some means other than that feature. Our discussion of the non-lateral approximant /r/ in the previous study suggests that /l/ and /r/ are distinguished by the presence of the feature [Coronal][1] in the feature geometry of /r/. This analysis predicts that the feature [Coronal] will be intimately linked to the acquisition of /l/ and /r/ and that these two segments will only be differentiated in a learner's grammar once [Coronal] has been acquired. In Chapter 2, these implications are explored and the acquisition of segmental representations, in general, is investigated; but, first, it will be necessary to review the two major approaches to child phonology in the literature in order to provide a context for the present study and to make explicit some of the assumptions that underlie this research.

Previous work on phonological development in children can generally be divided into two basic approaches. One approach is to view child language as something unique and (possibly) fundamentally different in nature from adult language (e.g., Ferguson & Farwell, 1975; Macken & Ferguson, 1983; Stoel-Gammon & Dunn, 1985). Of course, this position does not preclude the use of linguistic constructs to describe child language; however, the goal of this perspective is to characterize the child's knowledge (as evidenced by his or her productions) as a self-contained system, with little or no reference to the adult system being acquired. The use of linguistic constructs may help to capture the regularity of the child's phonological system, but it does not commit the researcher to any claims about the similarity between the child's system and the target language.

---

[1] Although Coronal is technically a node, it has traditionally been referred to as a feature. The difference between these two terms is not substantive for the purposes of this dissertation. Therefore, I will refer to Coronal interchangeably as a feature or as a node, depending on the context, in the remaining chapters.

The other main approach taken by phonological acquisition researchers is to treat child language as identical in nature to adult language (i.e., is a possible human language); differences between adult forms and the child's output (again, production) are attributed to phonological processes applied to underlying forms, which are themselves presumed to be identical to the forms posited in the adult grammar (e.g., Smith, 1973; Ingram, 1976; Macken, 1987). Since each stage of the child's phonological system (indicated by the child's productions) is assumed to be a possible human language (i.e., constrained by Universal Grammar), an effort is made to characterize the child's system using the tools of phonological theory. However, the child's phonemic system is not seen as an impoverished version of the target language and, although care is taken to show how the child's outputs relate to adult forms, this approach allows the child to posit phonological rules not attested in the target language; in other words, each stage of development is not necessarily a closer approximation to the adult system.

The research presented in this chapter shares certain assumptions with this second view of child phonology but departs significantly from it in treating the child's system as an immature state of the emerging (i.e., target) phonological system (rather than simply as a possible human language that is systematically related to the adult language). This view of child language follows directly from Pinker's (1984) *learnability condition* and *continuity assumption*. As the approach to child phonology and experimental research presented in this chapter is heavily influenced by these assumptions/external constraints, I will briefly explain how each constrains our theory of acquisition, paying particular attention to phoneme acquisition.

The learnability condition ensures that the child's linguistic system is viewed as an intermediate stage in the acquisition process; it also guarantees that our theory provides the child a way into the developmental state and explains how the adult system is ultimately attained. A system of representations cannot be posited to explain a given developmental stage, if one cannot show how those representations were acquired and how the proposed system evolves into a subsequent stage. Taking the adult system of segmental representations as the final state, the learnability condition sharply restricts the types of systems that can be proposed to characterize the child's knowledge of phonemic contrasts.

Whereas the learnability condition requires that the child's system is viewed as an intermediate stage in the acquisition process, the continuity assumption dictates that the child's intermediate stages are analyzed in terms of the adult system. In fact, (strictly taken) this constraint obliges researchers to regard the child's grammar as an impoverished version of the target system, rather than a unique system in its own right (regardless of whether that system can be shown to respect properties of Universal Grammar). The continuity assumption actually applies to three related aspects of the child's system: 1) the qualitative nature of the child's cognitive abilities, 2) the formal nature of the child's grammatical system, and 3) the realization of the child's grammar in comprehension and production. According to the continuity assumption, the null hypothesis is that, in each of these areas, the child and the adult are identical.

These external constraints on our theory of acquisition are a necessary departure point for any theoretically motivated approach to language acquisition: if our theory of the adult system is to inform our understanding and experimental investigations of first language acquisition, then we must assume that these two systems have *something* in common; if not, then it is pointless to use our theory of grammar to generate hypotheses about acquisition. By starting with an assumption that the child and adult system are identical in nature, we arrive at a constrained theory of acquisition that can be modified in a principled manner as the developmental data dictate.

Thus, in the research reported and discussed here, child phonology is not a unique phenomenon in and of itself but rather reflects an immature state of the emerging phonological system. In the context of segmental representations, this means that the representations posited by the child during the course of acquisition are assumed to be of the same type as the representations present in the adult system (i.e., feature geometric). Given that phonemes are distinguished from one another within a grammar by their internal structure, the acquisition of a phoneme involves the acquisition of that structure. Thus, in acquiring a phoneme (or phonemic opposition), the child must acquire the appropriate hierarchical organization of features to distinguish it from the rest of the phoneme inventory. However, it is not enough to demonstrate that children's productions can be analyzed in terms of segments whose internal structure respects the hierarchical relations encoded in Feature Geometry; we must determine how these feature-geometric representations are acquired given the theory of Feature Geometry presented in Chapter 1 and what we know about the acquisition process in general.

This chapter investigates how Universal Grammar (UG) and the primary linguistic data interact to create segmental representations in the learner's mental grammar. Our goal is to demonstrate that feature-geometric representations are, in fact, the end product of the acquisition process, and that the theory of Feature Geometry, and representations for laterals laid out in the previous chapter, make the right predictions regarding acquisition of these structures. It is hypothesized that the universal feature hierarchy encoded in UG actively constrains the acquisition of segmental representations by providing the material from which representations will be constructed and determining the order in which they will be acquired (Rice & Avery, 1991a, 1995). This theory of phoneme acquisition is tested with eighteen English children ranging in age from 1;3 to 2;4 who performed a forced choice picture selection task in which stimuli contained minimal pairs distinguished by a single feature in their underlying representations. The prediction is that children will discriminate phonemic contrasts that depend on less segmental structure before those that require more structure in an order consistent with the feature hierarchy. The results indicate that children indeed acquire the ability to contrast phonemic oppositions in a uniform order that is consistent with the hierarchical relations encoded in Feature Geometry.

2 The Role of Feature Geometry in the Development of Phonemic Contrasts

2.0 Introduction

Acquisition research within the framework of generative linguistic theory has generally restricted itself to the development of syntactic knowledge, while largely ignoring the role of Universal Grammar (UG) in the acquisition of phonological knowledge. However, the logical problem of language acquisition extends to the development of other areas of the grammar, including phonology. Despite imperfect primary linguistic data, children consistently achieve adult competence across the full range of subtle and complex properties of the phonological system of their language. Children must determine whether the language they are acquiring permits branching onsets, whether vowel length is distributionally predictable or phonologically contrastive, whether stress is sensitive to syllable weight, as well as many other options. At the segmental level, languages vary with respect to their phonetic (surface) inventories and their phonemic (underlying) inventories.

While the linguistic input provides the child with positive evidence of the variety of sounds that occur in the target language, the child must still determine whether those sounds are contrastive or whether they constitute allophones of a single underlying phoneme. The child must also develop phonemic representations for those segments in order to distinguish them in his or her grammar. It must, of course, be remembered that children have access only to the surface properties of the target language they are acquiring. Thus, any properties of the grammatical system ultimately acquired by a child that are not inducible from the input are presumed to follow from whatever biologically endowed apparatus the child brings to the acquisition task (i.e., UG). How the child determines which segments are contrastive in his or her language, how the representations for these phonemes are acquired and how UG guides this process are questions that are only now beginning to be addressed in light of current theoretical advances.

In an attempt to gain insight into how children acquire the phonemes of their language, researchers have often investigated the order in which phonemes are acquired. One of the first researchers to examine phoneme acquisition within a theoretical framework was Jakobson (1941). According to his theory of phoneme inventory development, there is a universal hierarchy of phonemic oppositions that determines the order in which all children will acquire specific contrasts, regardless of the language they are acquiring. A substantial body of research in child language development has accumulated which putatively refutes this hypothesis (Ferguson & Farwell, 1975; Kiparsky & Menn, 1977; Macken & Ferguson, 1983; Pye, Ingram & List, 1987). However, it must be recognized that Jakobson's claim actually contains two separable components: first, the hypothesis that a single hierarchical organization of phonemic oppositions determines the order in which all children acquire phonemic contrasts; and second, that the particular hierarchy of oppositions articulated by Jakobson, based on a synchronic study of a wide variety of phonemic inventories, is the correct hierarchy.

The evidence that has been used to argue against Jakobson's universalist position actually refutes the latter claim but remains orthogonal to the former. The vast majority of this literature cites pervasive variability across children from different language groups as well as across children within the same language group. Such variability, according to Jakobson's critics, is inconsistent with a model of phonological universals. This is true in the context of the strict hierarchy that Jakobson proposed. However, if we entertain an articulated hierarchy in which one opposition immediately dominates more than one other opposition, then Jakobson's model would predict a limited amount of variability yet maintain its universal application across individuals and across languages. Such a hierarchy would continue to make refutable predictions in that oppositions that are relatively low in the hierarchy will never be acquired before those oppositions that dominate them.

Recent developments in theoretical phonology have produced such an articulated hierarchy of oppositions, known more widely as Feature Geometry. According to this theory, distinctive features internal to the segment are arranged hierarchically (Clements, 1985; Sagey, 1986). This hierarchical structure captures the dependency relations that exist between features as well as defines the groups of features that pattern together in phonological processes. While this theoretical model is well supported by cross-linguistic phonological evidence, it has not yet been determined how such segment structure might be acquired by children.

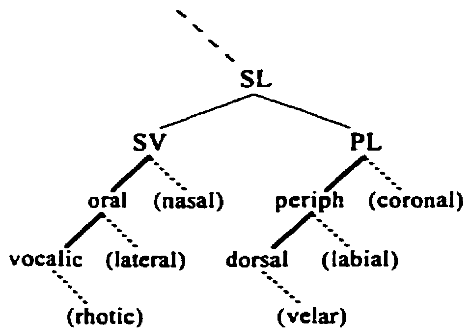
In this chapter we will reconsider Jakobson's (1941) theory of phonological universals in language development in light of current phonological theory and investigate what role Feature Geometry might play in the acquisition of segmental contrasts. We begin by reviewing several aspects of current phonological theory regarding the representation of phonemes (§2.1). We then consider how UG might interact with the linguistic input to create these phonemic representations. This discussion is followed by our own experimental study which measures children's abilities to discriminate phonemes in a comprehension task (§2.2). Based on the results of this experiment we infer a course of development for the acquisition of segmental representations (§2.3).

2.1 Theoretical background

2.1.1. Phonological theory and the representation of segments

The introduction of Feature Geometry as a theory of segmental representation has produced significant insights and advanced our understanding of segmental processes. In this framework the processes attested in natural languages are now believed to follow from the representation of the segment, in conjunction with a small set of principled operations, such as Spreading, Fusion and Delinking. Since the pioneering Feature Geometry models of Clements (1985) and Sagey (1986), there has been a proliferation of feature geometries, each modifying the structure and content of the original models in significant ways. A portion of one such model of feature geometry is given in (1) (Rice, 1992; Rice & Avery, 1991a, 1995).[1] We have chosen to focus on this geometry because Rice & Avery (1995) have used it to motivate a theoretical model for the development of phoneme inventories in first language acquisition.

(1)



	Node	Contrast represented
SL = Supralaryngeal	Sonorant Voice (SV)	sonorants vs. obstruents
SV = Sonorant Voice	Oral	oral vs. nasal sonorants
PL = Place	Vocalic	lateral sonorants vs. rhotics
	Peripheral	coronals vs. non-coronals
	Dorsal	labials vs. velars

With respect to the contents of a feature geometry, there is by no means a consensus among theoretical phonologists as to whether segments contain features specified for binary values or whether features are monovalent. For the purpose of our investigation, we will adopt the assumption held by many researchers that features are monovalent and that it is the mere *presence* of a feature in the representation of a segment that designates the active involvement of its corresponding articulator; likewise, the *absence* of a feature entails that the corresponding articulator is not active for a given segment (e.g., Anderson & Ewen, 1987; Avery & Rice, 1989; van der Hulst, 1989). For example, the voiceless segment /t/ will simply not contain the feature

[1] We list that portion of the geometry that Rice & Avery (1991a, 1995) themselves focus on. Their full geometry also contains an Air Flow (AF) Organizing node which would dominate the elements necessary for defining segments that contrast in terms of continuancy or stridency. Since they have not yet fully elaborated the internal organization of this branch of the geometry, we have excluded it from our investigation of the acquisition of segmental representations.

[voice] in its representation (that alone ensures that the vocal cords are not active for this segment), whereas the phoneme /d/ will be specified for the feature [voice].

Much of the work in phonological theory, and indeed linguistic theory in general, has been guided by the presumption that the lexicon contains only the idiosyncratic properties of lexical items and that predictable properties are derived (Chomsky & Halle, 1968). Thus, any redundant information (defined as that information that can be predicted or easily supplied by derivation, e.g., syllable structure) is absent from lexical representations. Within the present framework this principle is extended to segmental representations in what we will call Minimally Contrastive Underspecification (MCU). According to this position, a segmental representation contains only the information needed to contrast it from all other segments in the system. Any further specification will be provided by a system of phonetic implementation. Anderson and Ewen (1987) adopt similar assumptions within the framework of Dependency Phonology. MCU also resembles Clements' (1988b) degree-2 underspecification, termed Contrastive Specification, which attempts to remove redundancy from underlying representations without incurring the undesirable consequences of more radical models of underspecification (e.g., Archangeli, 1984, 1988; Pulleyblank, 1986).

However, MCU differs from Clements' model primarily in rejecting binary feature values. The features that are in parentheses in the model in (1) are generally not present in underlying representations of segments but instead are default phonetic implementations of underspecified representations. For example, many linguists believe that the coronal place of articulation is the universally unmarked place of articulation (e.g., Kean, 1975; papers in Paradis & Prunet, 1991). In terms of MCU, this means that coronal segments do not contain the feature [Coronal] in their underlying representation but receive it through the application of default rules (see Avery & Rice, 1989). A language will, however, require coronal segments to be specified for the feature [Coronal] if that language makes a phonemic contrast based on place of articulation within the class of coronal segments (e.g., alveolar vs. retroflex). Thus, which features a segment is specified for depends entirely on what that segment contrasts with in the particular inventory.

Importantly, Minimally Contrastive Underspecification refers to a property of representations, not to a process of the grammar. Other theories exist (most notably Radical Underspecification, Archangeli, 1984, 1988; Pulleyblank, 1986) in which underspecification is viewed as an active mechanism by which feature specifications for segments are operationally reduced to a theoretically defined minimum. In contrast, MCU maintains only that segmental representations contain all and only the features needed to distinguish them from the other segments in the system; moreover, features are introduced into a system only as they are required to differentiate segments in the inventory. There is no grammatical operation that "underspecifies" (or "de-specifies") segments. Rather, underspecified representations simply never become fully specified in the phonological component of the grammar.

Although Rice & Avery's geometry makes a few departures from the original models of Feature Geometry that are still widely accepted, there are certain properties that all models of Feature Geometry embody. Such properties include, for example, *dependency* and *constituency*. Dependency relations are represented structurally, such that the presence of a dependent node (or feature) in the representation of a segment entails the presence of its superordinate node in that representation. Constituency refers to the relation that holds among features that are dominated by a common node in the geometry. Since phonological processes manipulate constituents of segmental structure, the features of a constituent all pattern together in the phonological operations of a grammar.

With respect to the organization of features within the geometry, Rice and Avery make a distinction between what they call *organizing nodes* (referred to by Clements (1985) as "class nodes") and *content nodes* (or features). While the organizing nodes are assumed to be universal, languages differ as to which set of content nodes they utilize in segmental representations, depending on their phoneme inventory. Rice & Avery's geometry, like many others, contains separate "branches", each headed by an organizing node – Sonorant Voice (SV) and Place (PL). The Sonorant Voice Organizing Node organizes those features that pertain to sonorancy. The immediate dependent of the SV node is the feature [oral], which is used to contrast oral sonorants (/l/ and /r/) from nasals. Laterals and rhotics (/l/ vs. /r/) are then distinguished by the presence of the feature [vocalic], a dependent of [oral], in the representation of rhotics. Under the Place Node the feature [peripheral] distinguishes all peripheral (i.e., non-coronal) segments from coronal segments. This feature will be present in the representations of both /k/ and /p/ for example. The feature [Dorsal], below [peripheral], further distinguishes the dorsal peripheral segments (/k/ & /g/) from the non-dorsal peripheral segments (/p/ & /b/). Note that the notion of dependency entails that each of the segments, from /t/ to /k/ increases in what can be called "segmental complexity", such that the /t/ is the simplest segment containing the least amount of structure and the /k/ is the most complex segment containing the greatest amount of structure.

The hierarchy of Feature Geometry is assumed to be part of Universal Grammar (UG), so that while no one language exploits all of the features incorporated in the geometry, the geometry does represent all possible phoneme inventories of natural language. In this way the feature geometry is a substantive universal which provides information about what is possible in a language, not about what is impossible. Thus, rather than constrain a child's hypotheses, the feature geometry contained in UG comprises the material from which the child constructs the segmental representations in the phonological component of his or her grammar.

2.1.2. Acquisition of Segmental Representations

How do UG and the primary linguistic data interact to produce the phonological component of the adult grammar in general, and segmental representations in particular? Recent acquisition research has placed conditions on theories of grammar such that the proposed adult grammar must be learnable and that the initial state and end state of the learner's grammar may not differ substantively (e.g., learnability condition, continuity condition: Pinker, 1984). One of the consequences of these conditions is that children may not posit phonological representations or operations not present in the adult grammar of the language that they are acquiring (Ingram, 1989: 24). This is not to say that the immature system cannot differ from the adult system, only that the child's system cannot contain elements different from those found in the adult system. It is well known, of course, that child language frequently appears to have properties quite different from the adult system that the child is acquiring. For example, children acquiring English often exhibit a stage of linguistic development in which they appear to apply rules of vowel harmony or consonantal harmony (Menn, 1977, 1978; Vihman, 1978) despite the lack of evidence consistent with such a rule in the input. Rather than propose that such a child erroneously hypothesizes a rule that must then be subsequently "unlearned," we contend that some of what appears in child language to be rule-governed linguistic behaviour may actually reflect the default operation of principles and parameters of UG or result from their operation on impoverished phonological representations.

Thus, child language is not a unique phenomenon in and of itself but rather reflects an immature state of the emerging grammatical system. In the context of segmental representations, this means that the representations posited by the child during the course of acquisition will be of the same type as the representations present in the adult system (i.e., feature geometric). Viewing child language in this way provides a means of inferring the initial state of developing linguistic competence. Given that phonemes are distinguished from one another within a grammar by their internal structure, the acquisition of a phoneme involves the acquisition of that structure. Thus, in acquiring a phoneme (or phonemic opposition), the child must acquire the appropriate hierarchical organization of features to distinguish it from the rest of the phoneme inventory. A child will have acquired the phoneme inventory of his or her language once he or she has acquired the fully elaborated feature geometry of that language. How, then, does a child attain the adult feature geometry for the language he or she is acquiring?

The mechanism by which segmental representations are acquired

Based on Jakobson's (1941) hypothesis that a hierarchical organization of phonemic oppositions constrains how children acquire a phoneme inventory, we surmise two logical possibilities for the way in which Feature Geometry might govern this process. We have formulated these two possibilities as the Pruning Hypothesis and the Building Hypothesis, which are stated in (2).

(2) Pruning Hypothesis: Universal Grammar provides the child's emerging grammar with a fully elaborated feature geometry which subsequently retracts or is "pruned" where the input does not support phonological contrasts.

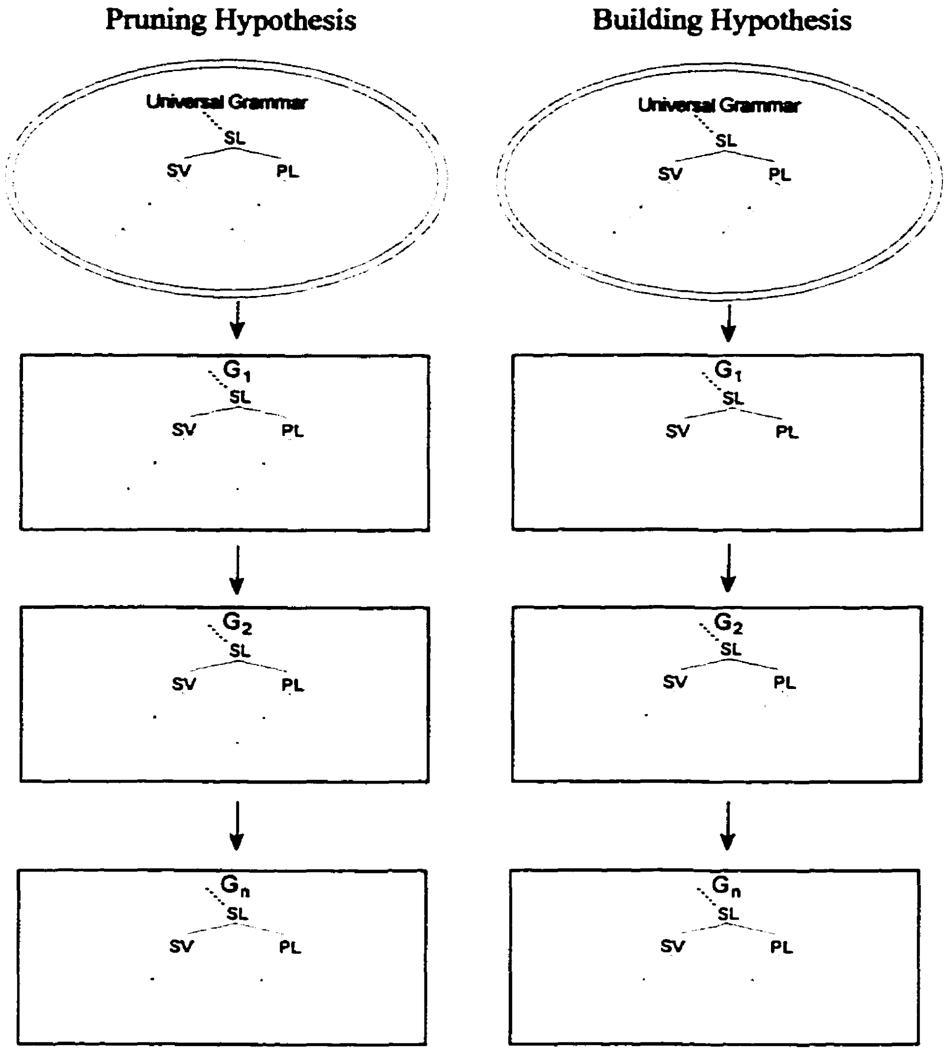
Building Hypothesis: Universal Grammar provides the child's emerging grammar with only minimal structure which is further elaborated based on the detection of phonemic contrasts present in the input.

While both of these hypotheses assume Feature Geometry to be a part of Universal Grammar and both ensure that the child will arrive at the adult grammar, the developmental stages that each of them predicts differ considerably. The Pruning Hypothesis parallels in many respects the strong continuity proposals of Hyams (1987) and Hyams & Wexler (1993) for the acquisition of syntax, while the Building Hypothesis resembles the proposals of structure building (Clahsen & Penke, 1992; Guilfoyle & Noonan, 1992; Radford, 1988).

The diagram in figure 1 illustrates the developmental sequences that each hypothesis entails. The ellipse at the top of the diagram represents the information about segmental structure that is present in Universal Grammar. Each of the subsequent boxes represents successive grammars hypothesized by the child, starting with G_1 (the initial state) and ending at G_n (the adult grammar).[2]

2 Note that we assume the position that properties of UG are transmitted to the emerging grammar, while UG remains intact (Haegeman, 1991; White, 1989a). This contrasts with the position that UG is, itself, shaped into the learner's grammar (Chomsky & Lasnik, 1993).

Figure 1



It is the intermediary grammars that contrast the two hypotheses: under the Pruning Hypothesis, the entire geometry is transferred from UG into the child's initial state grammar (G_1); once the child notices that his or her language does not contain a particular phonemic contrast, the content node that supports that contrast is "pruned". For example, when a child learning English realizes that velar segments do not contrast with pharyngeal segments, he or she will remove the structure that supports that contrast. This process continues until the child does not detect the absence of any other phonemic contrasts, at which point he or she will have arrived at the adult grammar (G_n). As shown in the diagram in figure 1, the feature geometry assumed by the child will get smaller and smaller until it coincides with the correct adult feature geometry for that language.

In requiring that children detect the absence of contrasts between sets of segments, the Pruning Hypothesis seems to rely on the child's sensitivity to what is not present or possible in the language (i.e., negative evidence). According to the Subset Principle (Berwick, 1985; Wexler & Manzini, 1987) children initially assume the most restrictive grammar consistent with the primary linguistic data available; only in response to positive evidence do children adopt a less restrictive grammar. It would appear, then, that the Pruning Hypothesis should be rejected solely on the grounds that it violates what we know about acquisition in general, namely that it is accomplished on the basis of positive evidence alone (Braine, 1971; Brown & Hanlon, 1970). This hypothesis could be maintained if the process were modified such that when a child observes that two phonemes are in complementary distribution or in free variation (i.e., positive evidence of their non-contrastive nature), he or she removes the structure that differentiates the two segments. However, all of the segments that are not contrastive in English (or any other language, for that matter) are not necessarily used in free variation or complementary distribution; some of them may simply not exist in the language. For example, velar segments do not contrast with pharyngeal segments in English, but neither are they in free variation; pharyngeal segments simply do not exist in English (at the lexical or phonetic level). Hence, this formulation of the Pruning Hypothesis cannot guarantee that a child will arrive at the correct adult grammar.

Under the Building Hypothesis, Universal Grammar supplies only minimal structure, namely that part of the feature geometry that is manipulated by every natural language, to the child's emerging grammar (G_1). This minimal structure is then elaborated by the child, guided by UG, based on the detection of phonemic contrasts in the input. Thus, for example, when an English child notices that labial segments, such as /p/, contrast with coronal segments, like /t/, he or she will posit the structural element that distinguishes these two places of articulation. This building process will continue until no more phonemic contrasts are detected, at which point the child will have arrived at the adult system. Thus, according to this hypothesis, the child's feature geometry expands, rather than recedes, until the adult geometry is acquired.

Both the Pruning and the Building hypotheses can be tested experimentally. The Pruning Hypothesis, which maintains that the fully-elaborated feature geometry is present in the child's grammar from the beginning, predicts that children should not exhibit consistent stages in the ability to discriminate phonemes in their language. In fact, very young children who have not yet begun to prune back their feature geometries should be able to phonologically discriminate non-native sounds in addition to those sounds contained in the inventory of the language they are acquiring. The Building Hypothesis, in contrast, predicts that children will only be able to phonologically discriminate those segments for which there is structure in their emerging feature geometries to differentiate. In other words, there should be discrete stages in a child's increasing ability to discriminate native sounds phonologically, as the child develops his or her phonological representations.[3]

Despite the objections raised above, the Pruning Hypothesis has received apparent empirical support from studies conducted on the ability of young infants to acoustically discriminate speech sounds. Research by Eimas, Siqueland, Jusczyk & Vigorito (1971), Kuhl (1979), Repp (1984) and Werker, Gilbert, Humphrey & Tees (1981) among others, has shown that infants as young as one month of age are able to acoustically discriminate all of the sounds of their native language, in much the same way as adult speakers of that language (i.e., categorically). This suggests that infants are able to sort acoustic variants of adult phonemes into some kind of category with relatively limited exposure to speech. Perhaps even more striking is the fact that these infants have been shown to possess the ability to discriminate many *non-native* speech sounds (i.e., sounds not present in the ambient language) (Aslin, Pisoni, Hennessy & Perry, 1981; Eilers, Oller & Gavin, 1978; Streeter, 1976; Trehub, 1976) suggesting

[3] A further prediction, which is not so easily tested experimentally, is that no child should phonologically discriminate sounds not present in the input language. This prediction follows from the fact that the feature geometry is elaborated based on the of detection of contrastive use only. If a child never hears two segments used contrastively, then there will be no impetus for him or her to add the relevant structure to his or her underlying feature geometry that would enable him or her to distinguish those two segments phonologically.

that the ability to acoustically discriminate speech sounds is, in fact, innate (Eimas et al., 1971). A decline in the ability to discriminate non-native contrasts is then observed between six and twelve months, while the ability to discriminate contrasts that are part of the native language remains (Tees & Werker, 1984; Werker & Lalonde, 1988; Werker & Polka, 1993; Werker & Tees, 1984a). Werker and her colleagues conclude from these findings that the decline in the ability to discriminate non-native phonetic contrasts occurs within the first year of life, just as the phonological system begins to develop.[4] Thus, it appears that while the ability to discriminate the potential phoneme contrasts of all languages is present at birth, specific linguistic experience is necessary in order to maintain this phonetic discrimination ability.

Although the developmental stages in the ability to acoustically discriminate speech sounds — from all sounds to only native sounds — appear to be consistent with precisely those predicted by the Pruning Hypothesis, it is important to consider exactly what type of knowledge these infant perception studies demonstrate. We will argue below, in the context of our own experimental results, that the findings from infant speech perception research do not constitute evidence of *phonological* development, but rather phonetic development. Consequently, it is doubtful that these findings are relevant to the distinction being made here between the pruning and building of phonological structure.

The Building Hypothesis appears to have received some empirical support from experimental research on the development of children's productive phonological abilities. Studies have frequently reported a gradual emergence of speech sounds in the segmental repertoires among young children. Early child speech consistently contains stops before fricatives, front segments before back segments and glides before liquids (Ferguson, 1978; Ingram, 1978, 1979). More recent acquisition research by Fee (1992) in addition to theoretical research by Rice & Avery (1991a, 1995) also support the Building Hypothesis. Rice & Avery (1991) have developed a theory of how phoneme

[4] Brown (to appear b) suggests that the very construction of segmental representations is responsible for the decline in infants' ability to discriminate non-native contrasts. The implications of this decline in perception for second language acquisition are also discussed.

inventories of languages are built up diachronically, while Fee (1992) has shown that the order in which children acquire the Hungarian vowel system reflects a systematic "building" process. The findings of Levelt (1992) for Dutch and Scullen (1992) for French also support this position. These spontaneous production data are not strong evidence for the Building Hypothesis, however, since they may reflect children's preferences rather than their full capabilities.

Research examining the comprehension abilities of young children has also provided evidence of their developing capacities to distinguish members of some contrasts before others (Barton, 1976; Edwards, 1974; Garnica, 1971, 1973; Shvachkin, 1948). These studies typically measure children's abilities to distinguish phonemic contrasts by means of a forced choice selection paradigm in which the child is asked to select an item (or picture) out of an array of items (two or three) that differ minimally from each other. Although these studies do assess the child's phonological knowledge — selection of the correct items requires that the child rely on an adequate internal phonemic representation to accurately perform the task — their results (in particular, the orders of acquisition they report) have been undermined by methodological, statistical and experimental design problems (see Barton, 1980, for a critique of this literature). Thus, the existing experimental data does not provide conclusive evidence in favor of either pruning or building; the experiment reported below tests these two hypotheses directly.

Feature geometric constraints on the mechanism of segmental acquisition

Before we examine the experimental evidence, however, we might refine the Pruning and Building Hypotheses, so that they make stronger and, therefore, more easily falsifiable predictions for acquisition. While both of these hypotheses maintain that retraction or elaboration of structure will occur monotonically, neither hypothesis dictates that these adjustments will occur in any particular order, either within a given child or across children. Yet, if we take seriously the issues of learnability and continuity, then the phonological structure posited in the child's grammar must reflect the organization of feature geometry in UG. Thus, by taking the fundamental properties of Feature Geometry into account, we can further constrain the way in which the emerging segmental representations are either retracted or elaborated. Although the principles of feature geometry apply equally to the Pruning and Building Hypotheses, for ease of exposition, we will only outline the implications of these principles for the Building Hypothesis (the converse of these statements will characterize the implications for the Pruning Hypothesis).

The notion of dependency requires that superordinate structure be posited in the child's grammar before dependent structure can be elaborated. This entails that there is a unique pathway of elaboration *within* a given Organizing Node: the child must add structure to an Organizing Node in his or her grammar in the order specified in UG (Rice & Avery, 1995). For example, the peripheral feature (which distinguishes /p/ from /t/) must be acquired before its dependent the Dorsal feature (which distinguishes /p/ from /k/) can be acquired. Thus, this notion of dependency strengthens the Building and Pruning Hypotheses by requiring that if there is an observable order to acquisition, it will be from unmarked to marked.[5] That is, simple segments will be acquired before structurally complex segments.

The notion of constituency, however, does allow for some variability in the order of acquisition across children. Although constituency requires that features within a given Organizing Node pattern together, Organizing Nodes remain autonomous with respect to one another. Given this autonomy, a child may selectively elaborate structure within any one of them, two of them or all of them simultaneously (Rice & Avery, 1995). Hence, if a child is first sensitive to the contrasting places of articulation of phonemes in the input (e.g., /t/ from /p/), he or she can begin to posit structure within the Place Node. Alternatively, another child may first detect differences in the

[5] When we order "unmarked" before "marked," there is no implicit claim that this order of acquisition *must* be observed, only that if an order is observed, it will be from unmarked to marked (White, 1989b). We say "observable order of acquisition" here because it may be the case that a child acquires the feature geometry before he or she is old enough to be tested by the means currently available. Alternatively, a child may detect the difference between two marked segments, say, /p/ and /k/ right away, building the relevant structure to differentiate these two segments. At the same time, however, the structure necessary to distinguish /t/ from /p/ and /t/ from /k/ would have been constructed in the process. In both of these cases, no order of acquisition would be directly observable.

sonorancy of phonemes in the input (e.g., /t/ from /n/) and elaborate the structure under the Sonorant Voice Node necessary to capture those contrasts. Thus, while the child is constrained as to the way in which he or she elaborates a particular Organizing Node, he or she has the freedom to elaborate structure underneath one Organizing Node before another, depending on what contrasts he or she initially detects in the input.

While the notions of dependency and constituency ensure that there will be a consistent, yet slightly variable, order of acquisition across children, the precise order in which phonemic contrasts will be acquired depends, of course, on the particular hierarchy (i.e., feature geometry) one assumes. For the purposes of our experimental study, we assume Rice & Avery's (1995) model of feature geometry, which has been motivated in part to account for the development of phonemic inventories. The principles of feature geometry, in conjunction with Rice & Avery's geometry, outlined above in (1), predict the order of phoneme acquisition given in (3).

(3)	<u>Sonorancy</u>	<u>Place</u>
	1 sonorants vs. obstruents	1 coronal vs. non-coronals
	2 nasal vs. oral sonorants	2 labials vs. velars
	3 laterals vs. rhotics	

Since the Sonorant Voice (SV) and Place (PL) Organizing nodes may be elaborated independently by the child, this model makes no predictions for the order of acquisition of sonorancy and place contrasts with respect to one another.

It should be noted that, because the Building Hypothesis has been developed from Jakobson's developmental theory of phonemic oppositions, it shares with that theory the characteristic of actually comprising two components: a theory of acquisition and a particular hierarchy of features. Although the two components are related, it is crucial to consider (and evaluate) them separately. The acquisition component consists of claims about what the acquisition process entails (i.e., that it is a *building* process) and how the process proceeds (i.e., on the basis of positive evidence and in a uniform order). These claims about acquisition, while wholly dependent on the *theory* of Feature Geometry, are independent of any particular geometry. The exact order of acquisition predicted, however, is dependent on a particular feature geometry. The following experimental study investigates both components of this proposal.

2.2 Experimental study

2.2.1. Rationale

As with any experimental acquisition study, how best to tap the learner's underlying competence is an important (and challenging) issue. There are two types of data that might be used to evaluate the proposal that Feature Geometry constrains the acquisition of phonemic contrasts. Production data would provide clear evidence that a child has acquired sufficient underlying structure to represent a given contrast when the child accurately and consistently produces the phonetic contrast that represents the underlying phonemic contrast. Unfortunately, a child might inaccurately produce a certain segment despite an adequate underlying phonological representation of that segment for reasons quite separate from the child's emerging phonology (e.g., the demands of the necessary articulatory gesture). We might mistakenly attribute less structure to a child's underlying phonological competence than he or she really has, based on the absence of a particular contrast from the child's production. By underestimating a child's phonological competence, data consistent with the Pruning Hypothesis might be erroneously interpreted to support the Building Hypothesis.

Comprehension data provide a similar opportunity to demonstrate that a child has developed sufficient underlying structure to support a phonemic contrast. If a child accurately and consistently distinguishes the two sounds of a phonemic contrast in comprehension, then we would argue that the child has the structure necessary to represent that contrast. Furthermore, if a child treats two segments as non-distinct, then we would argue that the child's underlying phonological representation is impoverished in such a way as to not contain adequate structure to distinguish the two sounds as separate phonemes.[6] Data from the comprehension abilities of children are taken to be a

---

[6] Of course, one way to increase the certainty that a child who fails to discriminate a phonemic contrast in comprehension truly lacks the relevant phonological structure would be to incorporate an

more accurate indication of the underlying competence than production abilities precisely because fewer peripheral performance systems intervene which might obscure the data collected. It must, of course, be recognized that the mechanisms of comprehension are no less part of linguistic performance than are mechanisms of production, and it is likely that the perceptual capacities of the learner must be sufficiently developed in order for phonological development to proceed. Nevertheless, with the two alternatives of comprehension and production tasks, we favor the comprehension task which we believe to have fewer potential confounds to obscure the child's underlying competence. An additional advantage of a comprehension task is that pre-verbal children may be tested; testing very young children is important in order to catch the earliest stages of phonological development, and, in fact, many of our own subjects were pre-verbal. Therefore, we have designed a comprehension task which assesses a child's ability to use his or her underlying phonological representations to distinguish phonemic contrasts present in the language that the child is acquiring.

2.2.2. Method

Subjects

Eighteen children who range in age from 1;3 years to 2;4 years (mean age = 1;11 years) participated in the study. Seventeen of the eighteen children attended a daycare facility where their participation in the study was made part of the unstructured play time, contingent upon parental permission. The eighteenth child was tested in a similar playtime environment at home. All children were equally familiar with the experimenters, who had been included in playtime activities for more than two weeks prior to the start of testing. This was done in order to reduce effects of shyness or anxiety in the actual testing session which might have depressed individual scores.

---

elicitation task into the paradigm. If a child could demonstrate the ability to accurately and consistently produce a particular contrast despite unsuccessful performance on the same contrast in our picture selection task, then our comprehension task could be shown to be underestimating the child's competence. Unfortunately, there is a limitation to the incorporation of a production task as many of the subjects included in this type of study (including our own) are pre-verbal. Nevertheless, our prediction would be that no child could produce a phonemic contrast consistently if he or she failed to discriminate that contrast perceptually.

Although eight of the children were simultaneously acquiring another language in addition to English, the phonological contrasts under investigation in this study are present in each of the other languages. In most cases the children had one parent whose native language was English and one whose native language was Chinese, French, Greek or Hindi. All of them are exposed uniquely to English in the daycare environment.

Materials

Each stimulus item contained an array of three color pictures (see Appendix A).[7] The names of all of the pictured items used in the test were single CV or CVC syllables. The names of two of the three items in an array constituted minimal pairs which differed only in their onset consonants.[8] The name of the third item differed maximally from the other two; the onset consonant differed in place of articulation and/or sonorancy from both of the other items in the array as did the vowel quality whenever possible. Each array was presented with a verbal cue that corresponded to one of the three pictures.

The experimental stimuli included six groups of ten picture arrays (for a total of sixty items). Each member of a group represented a phonemic contrast that can be captured by a node within the Rice & Avery geometry. For example, there were ten stimuli containing minimal pairs that contrasted velar from labial onsets, a contrast that requires a Dorsal Node in the underlying feature geometry. The contrasts examined in our study are listed in (4) along with the node needed to phonologically distinguish each contrast and an example of the contrast that was used in the test (see Appendix B for a complete list of stimuli).

---

[7] The materials used in stimuli preparation for the picture test reported here were adapted from the Bilingual Aphasia Test (Paradis & Libben, 1987).

[8] Care was taken to vary the vowel quality across the different minimal pairs as much as possible. Unfortunately, due to the practical limitations of finding minimal pairs that were both picturable and familiar to young children, there is an asymmetry in the vowel quality of our stimuli: the stimuli contain 19 minimal pairs that contain front vowels versus 11 pairs that contain a back vowel (the stimuli were relatively equal with respect to vowel height and tenseness). Although an analysis of our data with this asymmetry in mind indicates that it does not skew our results, future experimental designs should be careful to include equal numbers of the different vowel types.

(4) Experimental items:	<u>Node</u>	<u>*Example*</u>
sonorant vs. non-sonorant contrasts	SV	*nail—tail*
nasal sonorant vs. oral sonorant contrasts	Oral	*night—light*
vocalic sonorant vs. lateral sonorant contrasts	Vocalic	*rock—lock*
coronal vs. labial contrasts	Peripheral	*doll—ball*
coronal vs. velar contrasts	Peripheral	*tea—key*
labial vs. velar contrasts	Dorsal	*boat—goat*

The order of thirty test triads (i.e., minimal pairs) was pseudo-random. If any sequence of three identical contrasts occurred after randomly ordering the items, we separated the third item from the other two. The same order was then repeated for the second half of the test (hence, a total of sixty test items). Each occurrence of a particular triad was presented with a different verbal cue and each time in a different spatial arrangement. Every tenth stimulus item was followed by an array in which the verbal cue corresponded to the foil (i.e., neither member of the minimal pair). This was done to ensure that the child was attending to the task.

There were three possible responses for each stimulus item: correct (the correct picture), error (the contrast) or foil. In order to avoid any biases for position on the card, twenty items had the correct picture in the left position, twenty were in the center position and twenty in the right position. Likewise, the error picture and the foil each occurred twenty times in each of the three positions. The order of sixty picture arrays was reversed for a second test version to control for potential fatigue factors. Each test version was presented to nine of the eighteen subjects.

A training book was constructed which included every picture that appears in the experimental test. However, in this series each picture occurs in isolation (i.e., one to a page). The day before testing, each child went through a training session in which he or she saw each picture and heard the appropriate corresponding verbal cue. This was done in order to minimize the possibility that a child might make an error during testing simply due to his or her unfamiliarity with our particular illustration of an item.

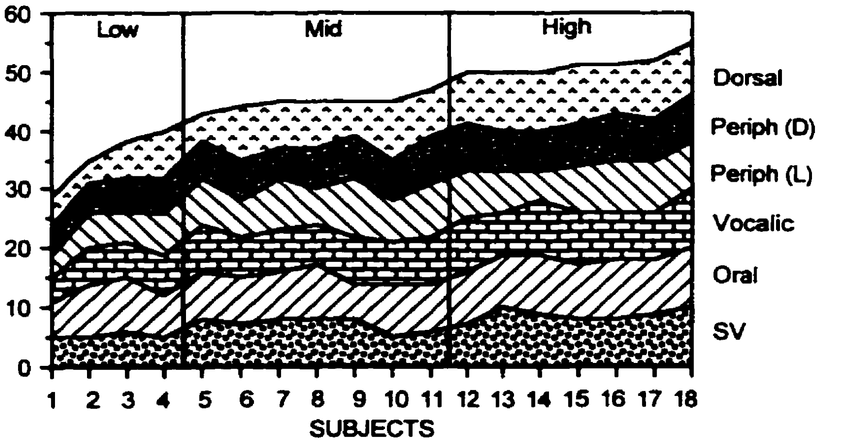
Procedure

Children completed a forced choice picture selection task. Each experimental session included two experimenters; one sat behind the child or held the child on his or her lap while the other was positioned in front of the child. The experimenter in front of the child engaged the child's attention and focused him or her on the book of stimulus cards positioned between them. The same experimenter ensured that the child scanned the entire array of pictures by pointing to each one and monitoring the child's gaze. The experimenter behind the child then produced the verbal cue and recorded the child's selection. The child heard "Point to the X" or "Show me the X" or "Where's the X" up to four times as needed by the child in order to make a selection. This set-up was used to minimize the chance that a child might detect visual cues that might aid in the identification of labials but not other segments, thus potentially inflating their performance on labials for reasons independent of the phonological system. Testing was stopped for the day if: (a) a child skipped three consecutive stimulus items, (b) a child showed other signs of decreased attention or (c) the child selected the wrong picture in an array in which the verbal cue corresponded to the foil. The sixty items were completed over the course of two to three consecutive days.

2.2.3. Results

The overall performance of each of the eighteen children is illustrated in figure 2. Each band across the graph corresponds to each contrast tested and is identified by the node underlying the contrast.

FIGURE 2
Overall performance on all contrasts



Dorsal = labial vs. velar contrast
Periph (D) = coronal vs. velar contrast
Periph (L) = coronal vs. labial contrast
Vocalic = lateral vs. vocalic sonorant contrast
Oral = nasal vs. oral sonorant contrast
SV = sonorant vs. obstruent contrast

A t-test revealed no significant difference in overall performance due to test version, t (16) = -.769, p >.05. This is true not only of overall performance but holds of performance on each individual contrast as well. A second t-test was performed on the performance of bilingual children versus unilingual children. Again, there was no significant difference for overall performance, t (16) = -1.612, p >.05, or performance on each contrast. The scores of the bilingual children and unilingual children are evenly distributed across the three performance levels: low, mid, high. We have, therefore, pooled the results from all of the children for our analysis of the data.

TABLE 1
Subject Performance Scores by Node[a, b]

Subject[c]	Age[d]	SONORANT			PLACE			TOTAL
		SV	Oral	Vocalic	Periph (L)	Periph (D)	Dorsal	
1	19	5	6	4	4	5	5	29
2*	15	5	**9**	6	6	5	4	35
3*	22	6	**9**	6	5	6	6	38
4	27	5	7	7	7	6	**8**	40
5	20	**8**	**8**	**8**	**8**	6	5	43
6*	26	7	**8**	7	6	7	**9**	44
7	21	**8**	**8**	7	**9**	5	**8**	45
8*	23	**8**	**9**	7	6	7	**8**	45
9	29	**8**	6	**8**	**10**	7	6	45
10*	28	5	**9**	7	7	7	**10**	45
11	23	6	**8**	**8**	**9**	**8**	**8**	47
12*	21	7	**9**	**9**	**8**	**8**	**9**	50
13*	21	**10**	**9**	7	7	7	**10**	50
14*	29	**9**	**10**	**9**	5	7	**10**	50
15*	20	**8**	**9**	**9**	**8**	7	**10**	51
16*	28	**8**	**10**	**8**	**9**	**8**	**8**	51
17*	21	**9**	**9**	**8**	**9**	7	**10**	52
18*	28	**10**	**10**	**10**	**8**	**8**	**9**	55

[a] Each score represents subjects' number of correct responses out of ten trials.
[b] Bolded scores indicate successful performance per our criterion (see text).
[c] Asterisks indicate subjects whose performance violates an order predicted by Rice & Avery's geometry.
[d] Ages are reported in months.

In addition, we found no significant correlation between overall performance and age, r = .416, p > .05. This can be seen in a comparison of age and performance totals in Table 1; for example, subject 4, who performed below criterion on all but one contrast, is twenty-seven months old, while subject 15, whose performance is near perfect, is only twenty months old.

Perfect performance on all test items would produce a score of 60. A score of 30 would reflect chance performance. Although the child selects one picture from an array of three, the actual choice is made between the minimal pair when the child is attending to the task. The foil is not an equally likely candidate when the child is making a selection in response to a verbal cue that shares no segments with the name of the foil. The graph in figure 2 illustrates the significant variation in overall performance across individual subjects.[9] The lowest scoring subjects performed at chance (29-39 out of 60) whereas the highest scoring subjects were at near perfect performance levels (50-55 out of 60).

The overall performance of each child does not provide sufficient information to determine whether certain contrasts are acquired before others. Furthermore, comparisons of groups would prove equally inadequate. Although members of a group have similar overall scores, their performance on individual contrasts might differ considerably. A child who performs perfectly on place contrasts but at chance on sonorant contrasts would attain an overall score very similar to a child who performs perfectly on sonorant contrasts but at chance on the place contrasts. Therefore, each child's performance on each individual contrast must be studied independently. Table 1 lists the scores for each child on each contrast measured.

We have set the score of 8 to be our criterion for successful performance on a particular contrast. Based on a standard binomial distribution, the probability that an individual could correctly choose between two equally likely choices 8 times out of 10 by random guessing is under 5% (Richmond, 1964).[10] We infer that if a child can

---

[9] Based on their overall performance, the children can be grouped into three distinct groups:

Subject groups:	Low	35.5/60 (group mean)	Significant difference in performance between all groups (p =.0001)
	Mid	44.9/60	
	High	51.3/60	

However, we refrain from making an argument based on the significant differences between the mean scores of these groups precisely because they are grouped based on their performance. Nevertheless, it is quite clear that the pool of 18 subjects does not constitute a homogeneous population with respect to the ability to discriminate all of the phonemes.

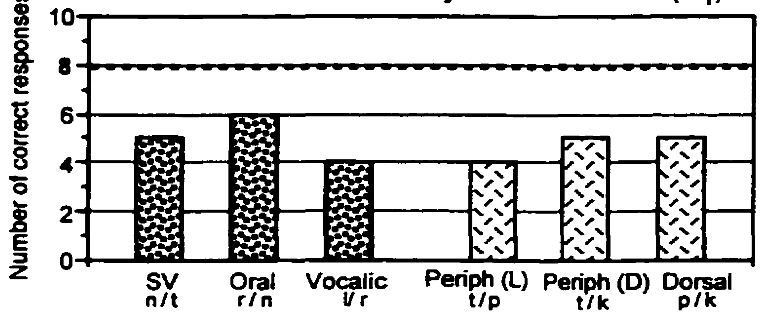
[10] Barton (1975) argues for the use of an even more stringent criterion (9/10) to establish that a child has successfully acquired a phonemic contrast. Based on a statistical formula for computing the likelihood that a particular score could be attained by random guessing, he claims that the probability that a child could make a correct choice 8 out of 10 trials actually falls above the 5% error range generally accepted for experimental studies in developmental psycholinguistics. We have conducted a

accurately distinguish minimal pairs that differ by only a single feature in a single syllable position 8 times out of 10, then he or she has sufficient underlying structure encoded in the emerging phonology to represent the contrast.[11]

Figure 3 contains the performance profile of an individual child from the Low group. Each bar corresponds to the child's performance on contrasts distinguished by a particular node within the Rice & Avery feature geometry. It is clear from this graph that the chance performance indicated by the overall score truly reflects the sum of chance performances on each contrast type tested. Only one subject in our study (subject 1) was tested at this early stage of phonological development.

FIGURE 3

Performance of child with only minimal structure (G_1)



In contrast to the performance illustrated in figure 3, children who scored at near perfect overall performance levels demonstrate abilities to distinguish phoneme contrasts that rely on a variety of different nodes within the feature geometry. The performance of an individual child depicted in figure 4 shows evidence that he or she has mastered all of the contrasts within the class of sonorants as well as the place contrasts. The asterisks in this figure and those that follow indicate a contrast on which

---

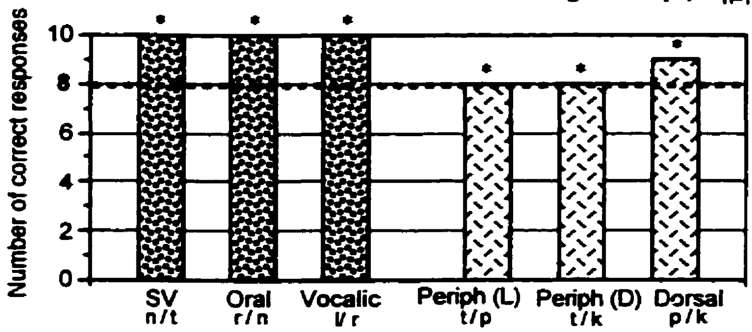
second complete data analysis using 9 out of 10 as the criterion for successful performance. Although several children are ascribed lesser discrimination abilities under the stricter performance criterion, there was no effect on the orders in which contrasts were acquired. Since this influence has no effect on our arguments, we report our results based on the less conservative criterion.

[11] Moreover, while it may be valid to establish a threshold at which we can infer with relative certainty that a phonemic contrast has been acquired, it is not equally valid to interpret any sub-threshold performance as indicative that a contrast has not been acquired. In particular, the difference between a score of 7 out of 10 and a score of 8 out of 10 might reflect the difference between a contrast that has not yet been acquired and one that has, but this cannot be determined conclusively based on scores that differ by a single error.

the child has met the criterion for successful performance. Five children exhibited near adult-like performance in which they reached criterion on at least five of the six contrasts tested (subjects 12, 15, 16, 17, 18). For those who did not reach criterion on all six contrasts, it was either the coronal/velar contrast or the sonorant/obstruent contrast that had not yet been mastered. These contrasts were consistently observed to develop later than others across children who demonstrated differential capabilities for different phonemic contrasts.

FIGURE 4

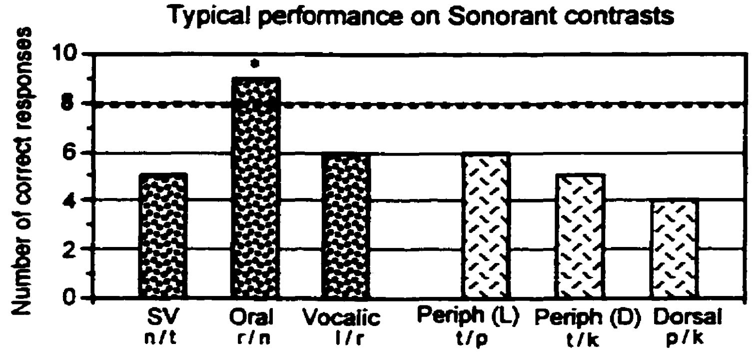
Performance of child with near-adult feature geometry (G_{n-1})



Although the performance profiles illustrated in figures 3 and 4 argue strongly in favor of the Building Hypothesis and against the Pruning Hypothesis, they are relatively uninformative with respect to assessing the claim that the order of acquisition follows from the organization of the feature geometry. The appropriate evidence that would speak to this issue can only come from children who have differential abilities in distinguishing phonemic contrasts of the input language (i.e., children in our Mid overall performance group). Despite the quantitative similarities in the overall scores of children in the Mid group, their performance on specific contrasts is qualitatively quite distinct. The performance of two children can be characterized by profiles like the one in figure 5 (subjects 2, 3). Such a child has clearly developed sufficient underlying structure to distinguish nasal from oral sonorants, yet he or she performed within the range predicted by chance on all other contrasts — including the sonorant–obstruent contrast which is subserved by Rice & Avery's SV node. The emergence of the nasal/oral sonorant contrast before the sonorant/obstruent contrast (an order that contradicts the predictions of the Rice & Avery geometry) occurs in every child who performs to criterion on only one of them (subjects 2, 3, 6, 10).

FIGURE 5

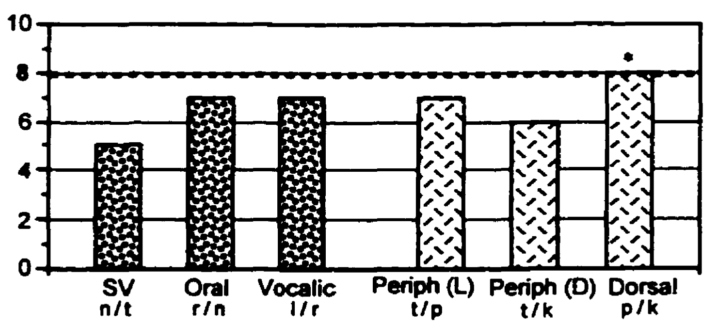
Typical performance on Sonorant contrasts



The same type of variable abilities is found among the place contrasts. The data in figure 6 illustrate the performance of a child who has acquired only enough structure to represent the labial/velar contrast. On every other contrast tested, this child performed within the range of chance. While only one child exhibits this ability to distinguish segments based on place of articulation but not based on sonorancy (subject 4), six children demonstrate the ability to discriminate labials from velars despite the inability to accurately distinguish either labials or velars from coronals (subjects 4, 6, 8, 10, 13, 14). Two other children accurately distinguished labials from coronals without reaching criterion on either of the other two place contrasts (subjects 5, 9), whereas none of the children ever reached criterion on the coronal/velar contrast without having attained that level of performance on all of the place contrasts.

FIGURE 6

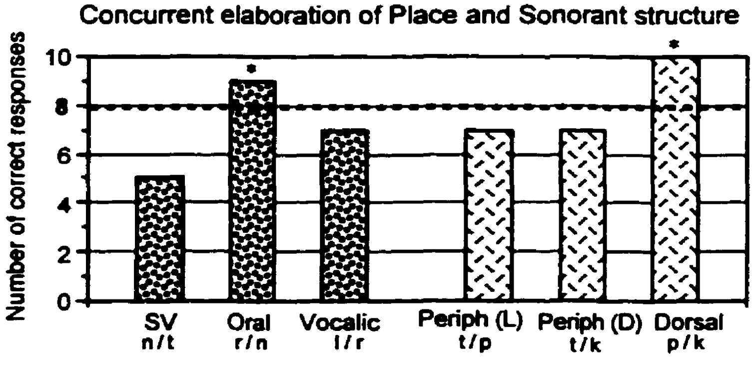
Typical performance on Place contrasts



The autonomy of Organizing Nodes provides two alternative learning paths, thereby allowing for variability across children in the elaboration of their phonemic inventories. One child might first elaborate structure within the Sonorant Voice Organizing Node (figure 5), while another child might begin by elaborating structure within the Place Organizing Node (figure 6). In fact, it is possible that a child might elaborate structure both within the Sonorant Voice node and within the Place node without having completely elaborated either one. Figure 7 illustrates precisely this type of behaviour. Within the place contrasts, such a child has acquired sufficient structure to underlie the labial/velar contrast but neither the labial/coronal nor the velar/coronal contrast. At the same time, the child has developed the structure necessary to distinguish nasal from oral sonorants but neither sonorants from obstruents nor oral sonorants from vocalic sonorants (i.e., laterals from rhotics). It is important to note that the first contrasts to emerge within each organizing node for this child are the same as those reported above (figures 5 and 6) for children who have elaborated structure under only one Organizing Node. Six children perform to criterion on only one or two of the contrasts within both Organizing Nodes demonstrating concurrent elaboration of structure within separate branches of the geometry (subjects 6, 7, 8, 9, 10, 13).

FIGURE 7

Concurrent elaboration of Place and Sonorant structure



In order to ensure that poor performance was not due to difficulty with particular stimulus items, we examined the performance on individual items. Difficulty with a particular item might be caused either by a child's inability to recognize the specific pictorial representation of the item or by a general lack of familiarity with the item (i.e., lack of a lexical representation). This difficulty should be reflected in the children's avoidance of the given item. There was no single item that was consistently avoided by all of the children (nor was there any item that all children, regardless of their phonological development, accurately chose).[12] Therefore, we conclude that our results are not skewed by bad stimulus items.

The training session was intended to familiarize each of the children with all of our stimuli as well as with our particular representations of those stimuli. The children's performance on some of the generally less familiar items suggests that this training session was indeed successful. For example, the item *bait*, whose picture consisted of a worm on a hook, was correctly chosen 72% of the time. Based on this type of performance, we feel confident that the training session provided adequate exposure to both the stimuli items and their pictorial representations.

To summarize, our data suggest the order of acquisition in (5).

(5) a) **SONORANTS:**

	no contrast (2)		*stage 0*
	oral vs. nasal sonorants (4)		*stage 1*
lateral vs. vocalic (2)		sonorant vs. obstruent (3)	*stage 2*
sonorant vs. obstruent	(6)	lateral vs. vocalic	*stage 3*

b) **PLACE:**

	no contrast (3)		*stage 0*
labial vs. velar (6)		labial vs. coronal (2)	*stage 1*
labial vs. velar labial vs. coronal	(3)	labial vs. coronal labial vs. velar	*stage 2*
	coronal vs. velar (4) labial vs. coronal labial vs. velar		*stage 3*

[12] The item *cape*, presented in contrast to *tape*, was only correctly chosen 22% of the time, suggesting that it might have been problematic for some children. However, we must keep in mind that this item involves a contrast between coronals and velars, which is one of the latest contrasts to be acquired. If children lack the relevant structure to differentiate two items, then their choice may be influenced by some extralinguistic factor (for instance, *tape* is surely more salient to a toddler than *cape*). In contrast, if the child does have the necessary phonological structure, the choice between two items should be made solely on the basis of phonological representations, cf. accurate performance on *bait* vs. *gate*.. Moreover, *cape* is a single item; if a child performed correctly on the remaining items within that contrast, acquisition of that contrast would have been attributed to him or her. However, this is not the case; we find that performance on the other items involving this contrast reflect random choices.

As the diagram in (5a) illustrates for sonorancy, two children failed to distinguish any of the phonemic oppositions tested. The first sonorant contrast to be acquired is oral vs. nasal sonorants (stage 1). At the next stage, children appear to have a choice. They may either add the lateral vs. vocalic contrast (2 children) or the sonorant vs. obstruent contrast (3 children). By stage 3, children have the ability to discriminate all of the phonemic oppositions we tested (6 children). The diagram in (5b) illustrates the pattern of development within the place contrasts. Note that no child was able to distinguish coronal segments from velar segments, suggesting that this contrast is one of the last to be acquired.

2.3 Discussion

The Pruning Hypothesis and the Building Hypothesis make the same assumptions about the nature of segmental information contained in UG and the shape of phonological representations in the adult grammar. However, they make opposite claims about the acquisition process and therefore opposite predictions about the interim states of the emerging phonology. According to the Pruning Hypothesis, children should distinguish native phonemic contrasts in their grammar from the very first stages of language development. Under the Building Hypothesis, we would expect that children will be unable to discriminate phonemic contrasts at the initial stages of language acquisition and that this ability will gradually improve over the course of acquisition.

We presented children with a forced choice picture selection task in which they were required to discriminate between minimal pairs in order to successfully perform the task. The children varied significantly in their abilities to accurately and consistently discriminate phonemic contrasts. Out of 60 trials, the number of correct choices ranged from 29 (48%) to 55 (92%). Among the subjects in this study, there was one who did not reach criterion on any of the contrasts measured, some who attained near perfect performance on all contrasts, and some who demonstrated abilities to distinguish some, though not all, of the contrasts. Based on these results, we conclude that our findings refute the Pruning Hypothesis and maintain that they support the Building Hypothesis. In fact, the finding that some children are capable of distinguishing all of the contrasts under examination while others are unable to discriminate any of them is enough to indicate that this capacity is not present from the initial stages of phonological development.

One might argue instead that these results indicate that our task may be too difficult for some of the children, thereby causing them to perform poorly for reasons quite independent of an impoverished grammatical competence. However, there is reason to believe that task difficulty is not confounding our results. With the exception of the poorest performer, the children in our study "found the task difficult" only for certain contrasts. It would be surprising to find that the difficulty in performing a task that is beyond a group's cognitive capacity is restricted to a specific set of stimuli. Furthermore, the inclusion of a foil in the selection array provides a means of detecting if a child is not "on task". The absence of foils chosen in the data we collected (even from the poorest performer) suggests that the children were, in fact, attending to, and therefore capable of performing, the task.

We considered the fact that the developmental sequences observed in the infant speech perception studies appear to reflect those predicted by the Pruning Hypothesis. By requiring infants simply to indicate detection of a change in the segmental make-up of the speech stream, the perception studies indicate that infants can hear the difference between the acoustic properties of different speech sounds. These studies further show that through continued exposure to a particular language, this acoustic (or phonetic) capacity gradually abates such that a child ultimately perceives only those contrasts that are part of the language being acquired. However, we would like to argue that the perceptual capacities required to accomplish the tasks used in these kinds of studies do not rely on phonological representations. Kuhl & Miller (1975) and Kuhl & Padden (1983) suggest that the ability to discriminate speech sounds is not unique to humans. Both chinchillas and macaques have been shown to be sensitive to phonetic boundaries in their acoustic discrimination of sounds, just as humans are (for the voiced-voiceless contrast and place contrasts, respectively). This seriously undermines the potential claim that the ability of infants to discriminate sounds is based on some internal phonological representation (i.e., grammatical competence).[13]

Moreover, according to the Pruning Hypothesis, infants begin with a fully elaborated system of *phonological* structure which is pruned back where the input does not contain contrasts that would require the full structure to support distinct representations. If the decline in infant speech perception capacities is taken as evidence to support this hypothesis, then individuals with developed phonological systems should never be able to perceive any contrast not present in the native language. However, it has been demonstrated that four-year old, as well as adult, native speakers of English retain the ability to discriminate certain non-native contrasts (for example, plosives vs. implosives, prevoiced vs. voiced stops, lateral vs. palatal clicks) despite their absence from the English inventory (Best, McRoberts & Sithole, 1988; Bond & Adamescu, 1979; Burnham, Earnshaw & Clark, 1986).

One role of segmental representations in the phonological component of the grammar is to distinguish lexical items from one another, thereby signaling differences in meaning. No such difference in meaning is ever supported in English by contrasts among the set of non-native sounds mentioned above. There is, therefore, no reason to suppose that the phonological system of English would include the segmental structure that underlies the representation of these non-native contrasts. Consequently, whatever mechanisms are available that permit the four-year old and adult English speakers to discriminate these non-native sounds, they do not depend upon phonological representations for those segments.[14] Similarly, there is no evidence that the performance of this same type of discrimination among infants would rely on properties

---

[13] We must, of course, be careful not to equate the infants' speech perception abilities with those of the non-humans: there are at least quantitative differences (e.g., number of trials required to reach criterion) between the infants' and non-human species' performance on these tasks. However, the ability of non-human species to discriminate speech sounds without the benefit of phonological representations suggests that infants too do not rely on phonological representations to discriminate speech sounds.

[14] See Brown (to appear a, b, c) for a theory which accounts for the adult speech perception of non-native contrasts in terms of the phonological *features* utilized in the L1 grammar (see also Wode, 1992, and Flege, 1992, for discussion of non-native speech perception in terms of L1 phonological categories).

of developing phonological representations.

Thus, the ability to acoustically discriminate two sounds is independent of the child's phonological representations (if any) of those two sounds. Some researchers believe that while speech processing may be specialized in adults, it is better explained by general auditory mechanisms in infants (Studdert-Kennedy, 1986). This latter position draws a clear distinction between infants' discrimination abilities and their emerging phonological systems. Hence, while we recognize the importance of the infant perception studies, we do not believe that their findings reflect the development of a *phonological* system. In contrast, by requiring children to distinguish between representations based on a single verbal cue, our experimental technique forces children to rely on their developing segmental representations and phonological properties of their developing grammar. With sufficient exposure to linguistic input, this phonological ability to make such distinctions expands to include all of the contrasts contained in the language being acquired.

We have argued that a child's inability to discriminate certain phonemic contrasts demonstrates that his or her grammar lacks the phonological structure necessary to differentiate those segments. That is, that a child's poor performance reflects an impoverished grammatical system. Is it possible that these results indicate an immature perceptual system rather than an undeveloped phonological system? If the child has not yet acquired the ability to map different acoustic signals to their respective categories, then two sounds might be perceived as a single sound. For example, both [p] and [k] would be perceived as /p/ (or as /k/, for that matter). Such a child would be unable to accurately discriminate [piyz] and [kiyz] in a comprehension task – even if the phonological representations for /p/ and /k/ were fully developed.[15] In this case, the errors would not be due to the phonological representation but rather the misperception (i.e., identification) of the cue. If, on the other hand, the child's phonological system is impoverished, he or she would be unable to discriminate phonemes in a comprehension

[15] We leave aside the paradox of how those representations would have ever been posited if [p] ad [k] were not distinguished by the perceptual system.

task because the segments are not differentiated in his or her grammar. In this case, there is no difficulty in correctly perceiving the sounds: [piyz] is perceived as /piyz/ and [kiyz] as /kiyz/. But once perceived, the cue cannot be matched with a unique lexical item.

What kind of empirical evidence might be used to decide between these two positions? The fact that the children in our study made errors is consistent with both accounts. However, the type of errors made by the children supports the claim that it is the children's phonological, not perceptual, system that is impoverished. When a child was unable to discriminate a minimal pair, the error usually occurred in both directions.[16] In other words, /kiyz/ was chosen for the cue [piyz] some percentage of the time, and /piyz/ was chosen for the cue [kiyz] some percentage of the time. This sort of random choice is what one would expect if the phonological representations are impoverished: the child hears a cue and must choose between two items that are not distinguished in his or her lexicon. The child's choice, then, should vary between the two items and may be influenced by some non-grammatical factor, such as recency or frequency effects. In contrast, if the deficiency were perceptual and two acoustic signals were mapped to the same category (making them indistinguishable), then errors should occur in only one direction. Both [piyz] and [kiyz] would be perceived as /piyz/ (or /kiyz/), hence the same item should be chosen, regardless of which cue is given. Thus, the types of errors made by the children support the claim that it is indeed the *phonological* system that is impoverished and which gradually develops. Moreover, because of the lack of correlation between age and performance, the developing ability to distinguish phonemic contrasts could not result from some maturationally determined mechanism of phonological acquisition. Rather, this finding is entirely consistent with the degree of variability permitted in a model with an articulated hierarchical organization of features that is elaborated based on an individual's experience with the

[16] There was a tendency for errors to be in the direction of coronals. That is, for labial and dorsal cues, the children were more likely to choose a coronal item than they were to choose a labial or dorsal item for a coronal cue. This means that coronals were correctly perceived as coronals more often (82% of the time) than labials were perceived as labials (76%) or dorsals as dorsals (67%). This tendency is consistent with Miller & Nicely's (1955) findings for adults and, therefore, probably does not reflect an undeveloped perceptual system.

primary linguistic data. The model advocated here would then predict a steady order of acquisition dictated by the feature hierarchy but would not be committed to a specific timetable for the acquisition of phonemic contrasts.

Of course, merely demonstrating that a child gradually acquires the ability to distinguish one segment from another in his or her grammar does not indicate that the internal structure of the segment constitutes a feature geometry (or that feature geometry plays a role in the acquisition of phonemes); segments could be distinguished from one another using SPE-style feature matrices. In order to establish that the segmental representations acquired by the child are, in fact, organized hierarchically, we must also demonstrate that the order in which the segments are acquired follows from this hierarchical organization.[17]

Given that children gradually develop the ability to discriminate the phonemic contrasts of the ambient language, we must determine whether there is a consistent order in which phoneme oppositions are added to their emerging phonological systems. In particular, we would like to ascertain whether this order is the one predicted by Rice & Avery's model of feature geometry. Our findings indicate that children's impoverished segmental representations do not respect the dependency relations encoded within Rice & Avery's feature geometry. In particular, one strong prediction — that no child will contrast labials from velars without also distinguishing coronals from labials and velars — is clearly refuted. Six children exhibited precisely this profile. Similarly, Rice & Avery's geometry predicts that no child will be able to distinguish nasal from oral sonorants without the concomitant ability to distinguish sonorants from obstruents. The performance of four of the children contradict this claim. In fact, every child who made at least one contrast within the class of sonorants contrasted nasal from oral sonorants. Thus, the particular feature geometry assumed in this experimental study is not supported by the acquisition data.

---

[17] Another way to show that young children's representations are organized hierarchically would be to examine their productions for phonological processes that fall out of the internal organization of the segments. The drawback to this tack, though, is that it requires children to be verbally expressive. However, at the earlier stages of phonological development, children frequently have not yet begun producing words.

This finding could be taken to refute Rice & Avery's geometry as an accurate characterization of the internal structure of segments or to indicate that Feature Geometry plays no role in the acquisition of phonemic contrasts. However, the children in our study did not vary wildly with respect to the sequences in which they acquired phonemic contrasts. In fact, among those children who have acquired some, but not all, of the contrasts under investigation, there was remarkable consistency as to which contrasts were acquired before others and which contrasts remained to be acquired. Therefore, it appears that the children's phonological development is indeed guided by some underlying principle of organization. We contend that this organization is the hierarchical organization of features contained in UG (i.e., feature geometry).

A plausible alternative explanation for the systematic order of acquisition might be that the observed order stems from acoustic or perceptual properties, rather than the phonological properties, of the respective segments. Under this account, the phonemic contrasts to be acquired first would be those that are acoustically or perceptually most dissimilar. In order to investigate this possibility, the acoustic properties of the contrasts under examination must be compared in order to determine whether those whose acoustic signals that are most distinct were indeed acquired first. Although we have not yet examined these acoustic properties systematically, it seems unlikely that properties of acoustic similarity or dissimilarity could explain our results: for example, the spectrograms of sonorants are quite distinct from those of obstruents; in fact, the difference between them is greater than that found between the spectrograms of nasal and oral sonorants. Yet, the children in our study consistently acquired the oral vs. nasal sonorant contrast before the sonorant vs. obstruent contrast.

It is, of course, possible that the acoustic characteristics of a segment may not correspond directly to its perception by the human auditory system. That is, acoustic dissimilarity may not be equivalent to perceptual dissimilarity. Therefore, in addition to considering acoustic properties, we must also consider the *perceptual* "distance" between the contrasts tested. One way researchers have measured the perceptual distinctiveness of sounds has been to examine the perceptual confusion of those sounds under varying noise conditions (Miller & Nicely, 1955). Those sounds that are perceptually more similar to one another are more easily confused with each other than those sounds that are perceptually more dissimilar. The perceptual-confusion data reported by Miller & Nicely (1955) show that the p/k contrast is the most fragile place contrast: even in the noise-free condition, /p/ and /k/ are confused significantly more with each other (11% of the time) than either the p/t or t/k contrasts (0.2% and 0.8%, respectively). The p/k contrast continues to be confused significantly more than the other two contrasts with the addition of differing levels of noise masking (21% vs. 3% and 4%, respectively; 30% vs. 12% and 14%, respectively). This indicates that /p/ and /k/ are perceptually more similar than either /t/ and /p/ or /t/ and /k/. Interestingly, the t/p and t/k contrasts are confused at roughly the same rate as each other for all noise conditions, suggesting that the perceptual distance between the members of each of these contrasts is equivalent.

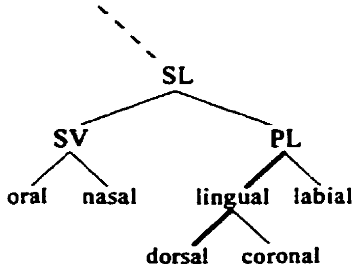
In terms of acquisition, these perceptual properties would predict that since the p/t and t/k contrasts are more robust than the p/k contrast, they should be acquired before the p/k contrast. Yet the children in our study exhibit the opposite order. Despite the fact that /p/ and /k/ are the most similar perceptually, it was the first place contrast to be discriminated phonologically by the children. Furthermore, since the p/t and t/k contrasts appear to be quite similar perceptually, we might expect them to be acquired at about the same time. However, we found that the p/t contrast was consistently acquired before the t/k contrast. Therefore, we can conclude that neither a perceptual nor an acoustic explanation is able to account for our findings. This, of course, does not deny any role for perceptual factors in the acquisition of phonemic contrasts, only that the order(s) of acquisition are not determined by these properties.[18]

Our results indicate that the development of phonemic contrasts is guided by a

[18] A reviewer has suggested that the uniformity we found across children might be due to the frequency distribution of particular contrasts in the input of native speaker English. This hypothesis predicts that children acquiring a language other than English might exhibit a different order of acquisition due to differences in the frequency distribution of certain contrasts. While we recognize that the frequency of particular contrasts in a child's input likely influences which contrasts are detected, we maintain that the underlying feature geometry encoded in UG constrains the order in which contrasts are acquired. These competing proposals could be tested empirically through a cross-linguistic study designed along the lines of the present investigation.

feature geometry contained in UG: the systematic order of acquisition, including a limited amount of variability, is consistent with principles of Feature Geometry. The particular feature geometry assumed in this study, however, is not supported by the acquisition data. Based on the findings illustrated in the diagram in (5), we propose the feature geometry in (6).[19]

(6)



The revised geometry differs from Rice & Avery's geometry in several significant ways. In contrast with the Rice & Avery proposal, we recognize [nasal] to be a feature in the geometry and not a default realization of the bare SV node (see also Piggott, 1992). Rather, a segment specified for SV but no further dependents of SV is realized as an oral sonorant (i.e., an approximant). Similarly, the modifications that we have introduced under the PL node represent a significant change from the model proposed by Rice and Avery. Rather than unite the labial and dorsal places of articulation under the feature [peripheral] as Rice and Avery do, we have introduced a feature, [lingual], which dominates the features [Dorsal] and [Coronal].

In an almost entirely articulation-based framework of phonological features, Browman & Goldstein (1986, 1989) propose a hierarchy of features that contains precisely this organization within place of articulation, and independent work in the study of consonant/vowel assimilation phenomena in Bantu also supports this structural organization of place contrasts (Zoll, 1993). In addition, Lass (1976) and Anderson &

[19] An anonymous reviewer has suggested that while our acquisition data appear to be consistent with the feature geometry proposed in (6) and inconsistent Rice & Avery's geometry, perhaps the geometry might change over time. Perhaps children begin with an internal organization of segment structure like the one in (6) but develop a geometry like the one proposed by Rice & Avery. We reject such restructuring on the grounds of continuity (Pinker, 1984). To permit such reorganization of internal segment structure would require some stipulation of constraints on restructuring in order to prevent the system from excessively overgenerating a range of possible structural organizations.

Ewen (1987) both recognize a structural constituent [lingual] in the organization of place features in internal segment structure. This structure under the PL node predicts that segments specified for the feature [Coronal] will pattern with segments specified for the feature [Dorsal] in phonological operations that manipulate the feature [lingual]. Such operations would effectively ignore labial segments which lack specification for the feature [lingual] (see Zoll, 1993). We believe that certain theoretical issues pertaining to the realization of segments underspecified for place features (i.e., as velar or as Coronal) can be accounted for with this geometry. We leave this as well as other interesting theoretical implications of this geometry for future research.

2.4 Conclusion

The results of our study indicate that children between the ages 1;3 and 2;4 gradually develop the ability to distinguish the phonemic oppositions of the language they are acquiring. Although the incremental addition of phonemic contrasts to the developing inventory appears to contradict the decremental decline in perceptual capacities reported in infant speech perception studies, we have argued that the different research paradigms actually measure the development of different levels of the linguistic system.

We would speculate that the changes observed in infant speech perception capacities are a necessary pre-cursor for the development of segmental representations in the emerging phonology. The initial (perhaps innate) capacity to perceive the difference between all of the speech sounds that occur in human languages predisposes the child to detect the differences in sound that could potentially signal a difference in meaning. By subsequently reducing the array of perceptual contrasts to include only those that will ultimately be manipulated by the emerging phonology of the particular language being acquired, the system permits increasing amounts of noise without loss of relevant information (i.e., linguistically meaningful differences). It is then the detection of contrastive use of different speech sounds that drives the process of segmental elaboration. The developmental sequence observed in the infant speech perception studies indicates how the system is adapted such that the child can ignore irrelevant variation in the speech stream yet remain sensitive to the sound differences that correspond to meaning differences.

In terms of how the acquisition process leads the child to ultimately arrive at the correct segmental representations for the language he or she is acquiring, our proposal makes the following claims. At the initial stage of phonological development, children will begin treating indiscriminately what are different segments in the adult system. Without sufficient structure to distinguish segments based on place of articulation, for example, children will confuse segments that differ in this contrast. This is not to say that they will not hear a difference between these segments, only that they will be unable to reliably use that difference in any linguistically meaningful way. Once the child detects that differences in place of articulation are used to signal differences in meaning, the organization of place contrasts present in UG constrains the process of encoding that difference in lexical representations. The initial stage of phonological development is characterized by treating members of separate phonemic categories in the adult language of the environment as members of a single category. Once the child detects that differences in sound correspond to differences in meaning, structure is added to the system in an order constrained by the hierarchical organization of phonological features encoded in UG. Based on the acquisition data collected in this study, we propose that the feature geometry in (6) accounts for that hierarchical organization.

If infants between the ages of 12 and 28 months have not yet mastered the full array of phonological oppositions in their language, how is it that they appear to understand so much of what is said to them? Why are they not walking (or crawling) around completely confused by so much indeterminacy in the primary linguistic input? The answer is that children can resolve such ambiguity in the same way that adults resolve the lexical ambiguity that pervades so much of natural language. For a child whose phonological development has not yet attained a level at which he or she can discriminate labials from velars, these segments contain the identical representation for place of articulation. Therefore, such a child's phonological forms of the lexical entries peas and keys would be identical, just as they are for pail and pale or bank and bank. Fortunately for the child, and for adults as well for that matter, there are typically syntactic, semantic and/or pragmatic factors available when decoding the speech signal. Our hypothetical child presumably avoids utter confusion upon hearing Have you eaten all your [$^{p}_{k}$i:z]? or Daddy left his [$^{p}_{k}$i:z] in the car by the same means that adult speakers use to resolve any lexical ambiguity.

As the development process continues, the child will nevertheless resolve the temporary ambiguity when further segmental structure has been elaborated. However, we would not expect the effect of the subsequent elaboration to be seen across the board. That is, individual lexical representations will still contain impoverished segmental structure. Therefore, although newly acquired forms will contain appropriately specified segmental representations, we would expect residual forms to persist for items that were present in the child's vocabulary prior to the elaboration of feature geometry. These forms will then be further specified on a lexical item by item basis. For this reason children will not exhibit discrete advancements from one stage of segmental elaboration to the next, although such stages will exist. The evidence of these stages will be blurred by the presence of the "archaic" forms held over from the previous stage of development.

Future research will be needed to elucidate with greater precision the relationship between the developing phonetic and phonological capacities of young children in the early stages of language acquisition. In particular, by testing the same set of children on a task that measures their ability to discriminate differences between two temporally adjacent acoustic cues and on a task that measures their ability to access a particular representation based on a verbal cue, we will begin to understand which characteristics of the capacities observed among young children reflect properties of phonetic and/or phonological development. Moreover, such testing conducted longitudinally would allow us to compare the receding phonetic capacities and the developing phonological capacities, which might speak to issues of continuity. Finally, if we were to conduct these types of experiments under varying noise conditions, we could more fully separate a grammatical explanation from a potential perceptual explanation of the developmental sequences that we found in the study reported here.

In reconsidering Jakobson's fifty year old theory of phonological universals in light of current phonological theory, we have (re-)discovered that the acquisition of a phoneme inventory in first language acquisition is guided by an articulated hierarchical organization of phonemic oppositions, which can be characterized in a theoretical model of Feature Geometry. We found a reliably consistent order in the development of phonemic contrasts across children that included only the amount of variability permitted by the nested structures of the feature hierarchy. Thus, contrary to a considerable body of literature has been published to refute Jakobson's universalist theory, we have demonstrated that by substituting a contemporary hierarchy of phonemic oppositions for the one Jakobson used in framing his theory, the substantive aspects of that theory, which pertain to the process of phoneme acquisition, are consistent with the acquisition facts. The putative counterevidence has been too tightly linked to the particular hierarchy of oppositions that Jakobson used to explain his theory and speak only indirectly to the theory of phonological universals.

One aspect of Jakobson's universalist theory which cannot be addressed with the data from our study is the claim that the same hierarchy of oppositions constrains the acquisition process regardless of what language a child acquires. Research in the acquisition of phonemic contrasts by children acquiring languages other than English will be necessary to investigate this part of the theory as well as whether the hierarchy that accounts for the development of the English phoneme inventory has universal application.

By demonstrating its active role in the acquisition of phonemic contrasts, we have established that Feature Geometry represents more than simply a formal tool for the description of segmental representations in order to account for phonological processes; as a theoretical characterization of innate properties of the developing human nervous system (i.e., part of UG), Feature Geometry constrains the process by which children acquire phonemic representations. Moreover, finding evidence for its role in acquisition lends external support to the theory of Feature Geometry, thereby strengthening and sharpening our conception of segmental representation. With continued investigation into the acquisition of phonology, we will increase our knowledge about the phonological properties of Universal Grammar, as well as broaden our understanding of the process of acquisition in general.

Preface to Chapter 3

Recall that the goal of generative linguistic theory is to develop a theory of grammar that accounts for observed cross-linguistic data and explains how acquiring that grammar is possible. With respect to our theory of segments, if feature-geometric representations are, in fact, an accurate characterization of the knowledge a speaker possesses about the phonemes in his or her language, then these representations must be learnable. The preceding chapter demonstrates that not only can feature-geometric representations be acquired given the necessary input, but that the Feature Geometry posited in the phonological component of Universal Grammar (UG) plays an active role in the acquisition of those representations. Thus, we have evidence from L1 acquisition for the representations proposed in Chapter 1 for /l/ and /r/, as well as for the theory of Feature Geometry itself.

The field of second language (L2) acquisition provides another opportunity to further test our theory of phoneme acquisition and our posited segmental structures, as well as to look for psycholinguistic evidence that these feature geometric representations correspond to psychological properties of a speaker's mind. To date, however, relatively little research has been done within the generative framework on the acquisition of L2 phonological knowledge and, in particular, on the acquisition of L2 phonemes. Those few studies that have investigated the acquisition of L2 segments using the tools of current Feature Geometry theory have not considered the role of UG in the L2 acquisition process (e.g., Flege, 1990, 1995; Hancin-Bhatt, 1994; Weinberger, 1990); moreover, these L2 phoneme studies have been conducted without reference to the L1 acquisition process, primarily because until now there has been no (experimentally supported) theory of how UG guides L1 phoneme acquisition.

According to the theory of phoneme acquisition developed in the previous chapter, UG initially provides the child's emerging grammar with that portion of the Feature Geometry which is utilized by all languages. This impoverished geometry is subsequently elaborated until the feature geometry of the target language (i.e., the adult geometry) is attained. The trigger for this elaboration of the child's geometry is his or her detection of contrastive use of segments in the ambient language: when the child encounters segments which cannot be accommodated by the existing representations in his or her mental grammar, the relevant structure is added to the emerging geometry. Thus, acquisition of phonemes is accomplished through the interaction of UG and the linguistic input; detection of a contrast in the input triggers elaboration of the child's geometry and, therefore, is a necessary condition for successful acquisition of segmental structure.

Having established the role of UG in the acquisition of L1 phonemes, the trigger for the elaboration of segmental structure and the process by which it proceeds, we can now consider the acquisition of L2 phonemes in the context of L1 acquisition (i.e., whether UG guides L2 acquisition in the same way that it does L1 acquisition). The following chapter investigates whether UG operates in the L2 acquisition of phonemes and formulates a theory of how UG interacts with the input and existing linguistic knowledge to construct novel segmental representations. The learnability condition and continuity assumptions presumed to constrain our theory of L1 acquisition are taken here to also constrain our theory of L2 acquisition. In the case of L2 acquisition, these constraints force us to assume that the same cognitive processes and grammatical mechanisms that ensure successful L1 acquisition also underlie L2 acquisition. Thus, our model of L1 phoneme acquisition does not merely function as a reference point for our L2 investigations, it serves as our null hypothesis as to how L2 phonemes are acquired. Our increased understanding of how successful phoneme acquisition is accomplished by the child will provide significant insight into why (the process of) L2 phoneme acquisition so often fails, and under what circumstances it is successful.

Studies of the operation of UG in L2 acquisition typically take the failure of L2 learners to acquire some aspect of the target language either as evidence that UG is not available in L2 acquisition (see e.g., Bley-Vroman, 1989; Clahsen & Muysken, 1986; Schachter, 1989) or that the L1 grammar impinges somehow on the L2 acquisition process, preventing the full operation of UG, which is otherwise accessible (see e.g., papers in Schwartz & Eubank, 1996; White, 1989). However, there are limitations to both of these perspectives. Denying UG a role in L2 acquisition certainly enables us to explicate the differences between L1 and L2 acquisition (the cases of acquisition failure, in particular); however, it makes it very difficult, if not impossible, to account for the many similarities between L1 and L2 acquisition (including successful acquisition). The position that UG is accessible but that the L1 grammar may prevent its operation in some circumstances sidesteps this problem; however, this approach fails, at least thus far, to fully explicate the relationship between the existing linguistic knowledge of a speaker, the operation of UG and the acquisition of a second linguistic system: it is still unknown which aspects of the L1 grammar impinge on the operation of UG, and we do not yet understand the mechanism by which the L1 grammar exerts this influence.

This chapter provides some answers to these questions by exploring the role of the L1 grammar in the acquisition of L2 phonemes. We know from our L1 acquisition research that two conditions are necessary for the successful acquisition of segmental structure: 1) UG must be operative, because it contains knowledge of the hierarchical organization of features that is not deducible from the input alone, and 2) the learner must be capable of perceiving the relevant segments as distinct in the input, since acquisition is only triggered by the learner's detection of the contrastive use of those segments in the ambient language. If either of these two conditions is not met, L2 learners will fail to acquire novel segmental representations and the non-native phonemes will not be differentiated in their interlanguage grammars. The research presented here demonstrates that those learners who fail to acquire a given non-native contrast (putative evidence that UG is inaccessible in L2 acquisition) also fail to perceive those segments as distinct sounds. Thus, the second criterion for successful acquisition – accurate perception of a contrast – is not met. Hence, the failure of these L2 learners to acquire the novel segmental representations cannot be taken as evidence that UG is inaccessible in L2 acquisition; the operation of UG can only be assessed in those cases where learners accurately perceive the L2 input. In fact, L2 learners who do perceive the non-native segments as distinct are shown to distinguish them phonologically in their interlanguage grammars, indicating that acquisition of L2 segmental representations is possible when the acquisition device receives the necessary input. The remaining question, and the focus of this chapter, is to ascertain what determines whether or not an L2 learner will accurately perceive a given non-native contrast.

Drawing together findings from infant speech perception research and the L1 phoneme acquisition data presented in the previous chapter, I develop a model of L1 phonological interference which explains how the elaboration of Feature Geometry in L1 acquisition restricts the sensitivity of the speech perception mechanism to non-native contrasts, and how this acquired phonological structure mediates speech perception in the mature speaker, effectively filtering out certain phonetic properties of the input. This model predicts that L2 learners whose native grammars lack a particular feature (e.g., [coronal]) will be unable to accurately perceive any non-native contrasts that are differentiated by that feature (e.g., /l–r/); conversely, the presence of the relevant feature in an L2 learner's native grammar will enable him or her to accurately perceive any two non-native segments distinguished by that feature. Acquisition of the segmental representations for non-native phonemes should follow in those cases where the contrastive nature of those segments can be detected in the input.

In order to test these predictions, Japanese and Chinese speakers' perception and acquisition of the English /l–r/, /b–v/ and /f–v/ contrasts is compared. The experimental results indicate that L2 learners do, in fact, accurately perceive only those non-native contrasts that are distinguished by a feature already present in their native grammar and that they successfully acquire only those contrasts which they perceive as distinct. Thus, we have evidence that properties of the native phonological system impinge on the perception of the L2 input; moreover, this research indicates that it is the featural inventory of the L1 which exerts this influence, determining which phonological properties of the L2 input the learner will be sensitive to and, hence, able to acquire.

3 The Role of the L1 Grammar in the L2 Acquisition of Segmental Structure

3.0 Introduction

There is still considerable debate as to whether the innate language faculty attributed to children (i.e., Universal Grammar, henceforth UG) continues to operate in the acquisition of a second language.[1] Experimental evidence adduced on either side of this issue is inconclusive and sometimes even contradictory (see for example White, 1989a, and Schachter, 1989). The acquisition of a second language is clearly *somehow* different from that of a first language: adult second language learners rarely (if ever) achieve the same native competence that children do learning their first language. This lack of success is often taken as evidence that UG does not operate in second language acquisition. But, perhaps there is another explanation. As White (1989a) points out, other factors, in addition to UG, are necessary for successful first language acquisition (e.g., sufficient input and various learning mechanisms); the same is true for second language acquisition. An intriguing line of research suggests that the failure of some L2 learners to attain a native-like competence is attributable to these other factors, rather than to the inaccessibility of UG (e.g., papers in Schwartz & Eubank, 1996).

One such factor that distinguishes second language acquisition from first language acquisition is the fact that the second language learner comes to the task of acquisition already knowing a language. Most current theories of L2 acquisition do, in fact, assume that the native language of the learner plays a role in acquisition; however,

---

[1] Although the term "second language acquisition" (SLA) technically refers to the acquisition of a second language by either an adult or a child, it is typically used to denote acquisition by post-pubescent learners. In this paper, we will only consider SLA by adults. The claims made here, however, can be extended to L2 learners of all ages.

what role the native language plays is less certain. Some researchers believe that aspects of the native language (e.g., parameter settings) are initially transferred from the native grammar to the interlanguage grammar but, given the appropriate input, will ultimately be readjusted to the correct L2 setting (e.g., Schwartz & Sprouse's *Full Acess/Full Transfer Hypothesis*, 1996; White's *Transfer Hypothesis*, 1988). On the other hand, others believe that the native language serves as a "surrogate" UG for the learner and that only those aspects of UG that are manifested in the native language will be acquired by the L2 learner (e.g., Bley-Vroman's *Fundamental Difference Hypothesis*, 1989). The position that the native language plays a role in the acquisition process is based primarily on research on the L2 acquisition of syntactic knowledge. Only recently have researchers begun to examine the role of the learner's native language in the acquisition of phonological knowledge (e.g., Archibald, 1993; Brown, 1993a, 1996, to appear b; Hancin-Bhatt, 1994a,b; Pater, 1993; Weinberger, 1990, 1994). This type of research shows that examination of the acquisition of phonological knowledge by second language learners can provide support for SLA theories which were initially based on syntactic data alone. Even more importantly, this phonological research has begun to define more precisely the role of the native language in the process of acquiring a second language.

Yet, despite the potential contributions of phonological research to the issues of UG-accessibility and L1 interference, it is tacitly assumed that the acquisition of phonology is fundamentally different from the acquisition of syntax and that it is, therefore, orthogonal to these SLA issues (Flynn & Manuel, 1991). The belief that the acquisition of syntax and the acquisition of phonology are qualitatively different is based not so much on experimental evidence as it is on general observations that L2 learners often have extraordinary difficulty mastering the pronunciation and intonation patterns of their second language, often retaining an accent when their syntactic knowledge of the L2 is quite native-like. This perceived disparity between the acquisition of syntactic knowledge and phonological knowledge (as well as a poor understanding of the role of UG in the L1 acquisition of phonology) has led some researchers to conclude that the acquisition of phonological knowledge by L2 learners is irrelevant to the question of whether or not UG operates in L2 acquisition (White, 1989a:42). However, recent work on the acquisition of prosodic structure by L2 learners suggests that learners do, in fact, acquire knowledge of phonological properties of the L2, including syllable structure (Broselow & Finer, 1991) and stress assignment (Archibald, 1993; Pater, 1993). Furthermore, experimental research by Brown & Matthews (1993, in press) demonstrates the active role of UG in the acquisition of phonological representations by children. Thus, the question of whether adult learners can acquire L2 phonology is indeed relevant to the issue of UG-accessibility in L2 acquisition, and research that addresses this question provides an opportunity to identify more accurately those aspects of the L1 grammar that affect L2 acquisition and perhaps even begin to characterize the psychological mechanisms by which this influence is exerted.

This paper will attempt to shed light on the issues of UG-accessibility and the role of the L1 by exploring the following research questions from the perspective of phonological acquisition:

1. How do we account for the failure of L2 learners to acquire (phonological) properties of the target language provided by UG?
2. Which aspect(s) of the L1 grammar impinges upon L2 (phonological) acquisition?
3. When the necessary conditions for successful acquisition are met, can L2 learners acquire novel (phonological) representations?

We will take as our starting point the hypothesis that differences between L1 and L2 acquisition (e.g., course of development, ultimate attainment) can be attributed to the L1 (rather than, say, to different underlying cognitive processes) (Schwartz, 1996). However, the position that it is the L1 grammar which prevents successful acquisition of L2 (phonological) knowledge is not particularly illuminating since we are still left wondering how exactly the L1 constrains L2 acquisition. Does the L1 grammar serve as an inventory of linguistic components from which the interlanguage grammar is constructed, constraining acquisition by restricting the *elements* of acquisition? Or, does the L1 grammar play a more dynamic role, perhaps intervening between the input and the acquisition device, and thereby disrupting the *process* of acquisition itself, by restricting which properties of the input the learner will be sensitive to? Thus, in addition to demonstrating that properties of the interlanguage can be traced to the L1 grammar, we must also explain why this transference takes place and attempt to identify the mechanisms by which it occurs.

By investigating the acquisition of non-native phonemes by L2 learners, I hope to provide some insight into the role that the L1 grammar plays in L2 acquisition: we will see that the L1 grammar intervenes between the L2 input and the acquisition device, effectively delimiting the properties of the input that the learner will be sensitive to and, hence, able to acquire. Researchers have long observed that learners are sensitive to different aspects of the input at different times. Therefore, it is necessary to make a distinction between *input* and *intake* (Corder, 1967). Input refers to the ambient language that the child is surrounded by every day, whereas intake refers to that portion of the input to which the learner is actually sensitive (i.e., the real input to the language acquisition device). Although input is required for language acquisition, the specific intake will determine the developmental stages of the grammar. This distinction between input and intake will prove to be crucial to our understanding of the acquisition of L2 phonology.

In the case of phoneme acquisition, (at least) two conditions are necessary for successful acquisition to occur: the operation of UG + the learner's detection of a phonemic contrast in the input. Thus, accurate perception of the L2 acoustic signal is crucial for phonological acquisition; if a learner does not perceive a contrast in the input, then (whether or not UG is operative) acquisition will not be triggered. Therefore, we cannot investigate learners' acquisition of phonemic representations without also examining their speech perception capacities. It is well known that one's native language experience has a profound influence on the ability to perceive speech sounds, both as an infant and as an adult, making non-native contrasts difficult, if not impossible, to perceive. Yet, to date, we do not have an adequate explanation for how this influence of the native grammar on speech perception originates or what impact this will have on phonological acquisition.

This paper develops a model of phonological interference that explains how the intake to the acquisition device is determined by phonological properties of the L1 grammar; in particular it is proposed that the monotonic acquisition of phonological structure by young children restricts their sensitivity to particular non-native contrasts and that the continued operation of this existing phonological structure in adult speech perception constrains which non-native contrasts adult learners will be sensitive to in the L2 input and, therefore, able to acquire. By bringing together research on phonological acquisition and infant and adult speech perception research, we arrive at a unified account of L1 phonological interference – from when and why the emerging phonological system begins to constrain perception to how this phonological system mediates adult speech perception, and the ramifications of the interrelation between the phonological system and speech perception for the acquisition of phonological knowledge by L2 learners.

In order to test the predictions of this model of phonological interference, the acquisition of segmental representations by Japanese and Chinese learners of English is examined; two separate experimental studies investigate whether these L2 learners can learn to represent the non-native /l–r/, /b–v/ and /f–v/ contrasts in their interlanguage grammars. These experimental studies demonstrate how the L1 phonological system may block accurate perception of the input, thereby preventing the acquisition of novel segmental representations; in fact, in those instances when L2 learners fail to acquire an L2 phonemic contrast, the second condition necessary for successful acquisition (detection of the contrast) is not met. On the basis of these data, it will be argued that the inability to acquire new phonological representations is not due to the inaccessibility of UG, but rather to insufficient intake to the acquisition device. These two studies also establish the circumstances in which the native grammar actually facilitates perception of non-native contrasts, demonstrating that when the acquisition device receives sufficient intake, novel segmental representations can be successfully acquired.

The remainder of the paper will be organized as follows: we will begin by reviewing the relevant aspects of the theory of phoneme representation assumed in this study as a way of examining exactly what type of structures L2 learners must acquire (§3.1). This will be followed by a review of the development of the native phonological system, which will allow us to establish the conditions necessary for successful L2 phoneme acquisition. The following section (§3.2) presents a theory of phonological interference based on a comparison of the findings from infant speech perception research and L1 phonological acquisition research. The model of speech perception development outlined here establishes how the influence of the L1 grammar in adult perception originates and also identifies the mechanism by which the L2 input is mapped to existing phonological categories. After the implications of this theory for L2 phoneme acquisition are laid out, the results of two experimental studies will be reported and discussed. Experiment 1 (§3.3) investigates the acquisition of /l/ and /r/ by speakers of Japanese and speakers of Chinese, while experiment 2 (§3.4) compares the acquisition of the /l–r/, /b–v/ and /f–v/ contrasts by Japanese speakers. The paper concludes by considering some of the implications of these results for the theory of phonological interference as well as our theory of second language acquisition.

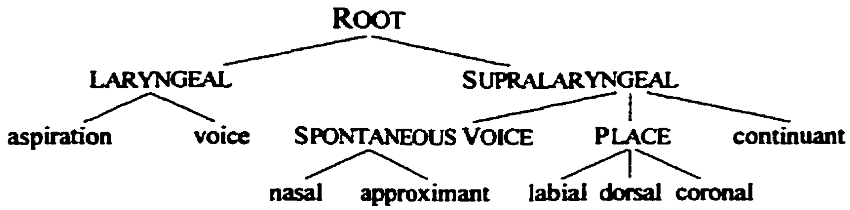
3.1 Knowledge of Phonemic Contrasts

In order to determine whether adult learners can acquire this type of knowledge about their L2, we need to first consider how the knowledge of phonemic contrasts is represented by a native speaker and how that knowledge is acquired in L1 acquisition. Once we establish what acquisition of segmental representations entails and how this process occurs, we will be in a better position to understand why this process fails so often in L2 acquisition. Part of the phonological knowledge a speaker possesses about his or her language is what phonemes are contained in its inventory. But, as Hockett (1958:24) points out, a phonological system is "not so much a set of sounds as it is a network of differences between sounds." Thus, a native speaker's knowledge of the phonemes in his or her language is tantamount to knowing which segments are contrastive in that language.

3.1.1. Representation of Phonemes

The theory of segmental representation known as Feature Geometry maintains that phonemes consist of distinctive features that are organized into hierarchical constituents (Clements, 1985; Sagey, 1986). Although there is still debate as to the precise set of features and their organization, phonologists agree that phonemes have an internal structure. The model of Feature Geometry that will be assumed in this paper is given in (1).

(1)



This model integrates insights of models proposed by Clements & Hume (1994), Piggott (1992) and Rice & Avery (1991b); however, the arguments and findings presented here do not hinge on the correctness of this particular model. This Feature Geometry is contained in the phonological component of UG. Like a syntactic principle or parameter, this geometry constrains the acquisition process and provides the learner with information about what phonemic oppositions are possible in natural languages. Thus, while no one language manipulates all components of this universal Feature Geometry, every phoneme in the world's languages can be represented in terms of the features and structural relations present in this geometry.

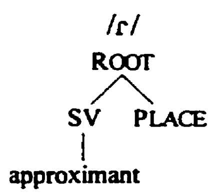
Each phoneme has a unique structural representation (i.e., feature geometry) that distinguishes it from other segments in an inventory. The representation of a given segment will be a subset of the entire geometry given in (1); its precise representation will depend on which segments it contrasts with in the particular inventory. For example in English, where the lateral approximant /l/ and central approximant /r/ are contrastive, /l/ may be represented as in (2a) while /r/ may be represented as in (2b), omitting unnecessary structure for the moment (Brown, 1993b, 1995).[2]

(2)

Both of these segments have the same manner of articulation (i.e., approximant), yet they differ in their specification for *place* features: /r/ contains the feature [Coronal] whereas /l/ is represented by a bare Place node. The fact that these segments have different representations reflects the fact that they are contrastive phonemes in English.

In contrast, when these two segments are not contrastive in a language, they will not have distinctive representations. For example, in Japanese, /l/ and /ɾ/ are allophones of a single phoneme, so there will be only one underlying representation for these two surface phones. This is given in (3).

(3)



The phonetic realization of this segment as an /l/ or an /ɾ/ (which, unlike the English /r/, is a flap) varies more or less freely (Vance, 1987).

---

[2] Brown (1995) is included in this disseration as Chapter 1. The representations for /l/ and /r/ assumed here differ somewhat from standard representations. The phonological feature [lateral] is generally assumed to distinguish laterals from non-laterals, in this case /l/ from /r/. However, Brown (1995) argues that [lateral] is not tenable as a phonological feature and that the contrast between /l/ and /r/ is best captured in terms of Place features (see reference for specific theoretical motivation and empirical evidence to support this claim; see also Piggott, 1993 and Spencer, 1984 for this view). Importantly, the representations for /l/ and /r/ given in (2) provide an explanation for differential acquisition effects due to a speaker's L1 grammar, which is not expected given the standard view of laterals. This will be discussed in greater detail in footnote 30.

A speaker's knowledge of which sounds in his or her language are contrastive is, thus, represented by distinctive segmental representations in the speaker's mental grammar. The following section examines how these representations are acquired in first language acquisition, where we will identify the factors necessary for successful (L2) acquisition.

3.1.2. Acquisition of Phonemes

Since segments are distinguished in a grammar by their particular feature geometries, acquisition of a phoneme contrast involves the acquisition of the relevant structure that differentiates the two segments (Rice & Avery, 1995, based on Jakobson, 1941). The learner is able to contrast the two segments in his or her grammar once that phonological structure has been acquired. Brown & Matthews (1993, in press) demonstrate experimentally (using a lexical comprehension task) that children's ability to differentiate phonemes phonologically develops gradually over time (see also Barton, 1980; Edwards, 1974; Garnica, 1973; Shvachkin, 1948 for similar studies).[3] Based on these results, Brown & Matthews argue that UG provides the child's emerging grammar with a minimal amount of segmental structure (in fact, only those portions of the feature geometry that are universal) which is subsequently expanded over the course of acquisition until the adult feature geometry for the particular language is attained. In other words, the child "builds" the developing geometry when phonemic contrasts are required.[4]

Brown & Matthews found that not only do children acquire phoneme contrasts gradually but they do so in a systematic order that is consistent across children. This

---

[3] Although it is generally recognized that children's production of different phonemes gradually develops over time, it is usually assumed that the underlying representation of these segments (i.e., representation of lexical items) is adult-like. Studies of child phonology, then, focus primarily on providing an explanation for the discrepancy between the adult-like underlying form and the child's production (e.g., Smith, 1973). Brown & Matthews (1993, in press) , in addition to the other studies cited in the text, is one of the few studies that demonstrates that children's underlying representations are impoverished (i.e., not adult-like) and that their developing knowledge of phonemes involves the elaboration of these underlying phonological representations.

[4] This position contrasts with the possible position that the child "prunes" the universal geometry where phonemic contrasts are not required by his or her language.

systematic order of acquisition results from the hypothesized acquisition process and the nature of Feature Geometry itself. UG, in addition to providing the initial structure, also guides the acquisition process: the child will only elaborate the feature geometry in his or her grammar in ways that are consistent with the hierarchical organization of feature geometry in UG. Thus, the particular dependency and constituency relations that are encoded in the feature geometry in UG will be respected in the geometry posited by the child. For example, the presence of a dependent feature in a representation entails the presence of that feature's superordinate node; by extension, superordinate structure must be posited in the child's feature geometry before dependent structure can be elaborated. As a result, children will phonologically distinguish those segments that require less structure to differentiate before distinguishing those segments that require highly articulated structure.

What is important for our purposes here is that this step-by-step elaboration of the child's feature geometry is driven by the child's detection of contrastive use of segments in the input (Rice & Avery, 1995). Once a child notices that two segments are used contrastively (i.e., are distinct phonemes), the phonological structure that differentiates the two segments is added to his or her grammar. If the child never encounters contrastive use of two segments (because, for example, they are allophones of a single phoneme in that language), the structure that differentiates them will never be posited. Therefore, input drives the acquisition of phonemic contrasts. However, the mere *presence* of a contrast in the input (while necessary) is not sufficient to trigger the further elaboration of the internal geometry; the child must *detect* the contrast in the input. In other words, the contrast must be part of the learner's intake in order to force the addition of segmental structure to the emerging grammar.[5]

In order for a learner to detect that two sounds are used contrastively, the learner must be able to discriminate the two sounds acoustically. Given that a child may be

[5] In fact, the child's current grammar seems to determine which contrasts he or she will be sensitive to. In particular, it appears that only once a child has elaborated a sufficient amount of the geometry will he or she be sensitive to contrasts that require additional structure. For example, a child will only be sensitive to the fact that nasals and liquids contrast (a contrast which requires a feature that is a dependent of the SV Node) once she or he has detected that the larger classes of sonorants and obstruents contrast with one another (a contrast which depends on the SV Node itself).

born into any language environment, it is imperative that he or she be equipped with adequate machinery to perceive the whole range of possible phonetic contrasts. Researchers have, in fact, demonstrated that infants as young as one month old are able to acoustically discriminate not only the sounds of the ambient language but non-native contrasts as well (Eilers, Gavin & Oller, 1982; Eimas, Siqueland, Jusczyk & Vigorito, 1971; Streeter, 1976; see Mehler, 1985 for a review), suggesting that the ability to acoustically discriminate the entire range of speech sounds manipulated by natural languages is innate (cf. Burnham, 1986). This predisposition ensures that the child will be able to detect contrastive use of any two sounds in the input and, therefore, that he or she will successfully acquire the phonological structure required to differentiate native phonemes in his or her mental grammar.

3.2 A Theory of Phonological Interference

Since the detection of contrasts in the input is crucial for the acquisition of segmental representations, we need to consider how this capacity changes (if at all) as the child develops. In particular, whether or not an L2 learner is able to perceive a non-native contrast (which is independent of the availability of UG) will be crucial in determining if he or she will be able to construct the phonological representations necessary to distinguish the two segments phonologically. It is well established that the ability to acoustically discriminate non-native contrasts decreases rapidly with exposure to a specific language, until the child is only able to acoustically discriminate those contrasts present in the language being learned – with the exception of certain non-native contrasts to be discussed below (see Werker & Polka, 1993, for a review of studies that establish this observation). This results in the categorical perception of speech, by which speakers of a language are able to easily distinguish members of different native phonemic categories and relatively unable to distinguish members of the same native phonemic category (Repp, 1984). In the following section, I examine some findings from infant speech perception research and suggest a causal link between the development of a learner's feature geometry and the subsequent decline in perceptual capabilities. Establishing such a link will have important consequences for the acquisition of L2 phonology.

3.2.1. Developmental Changes in the Perception of Contrasts

Some of the most interesting work in infant speech perception has been conducted by Janet Werker and her colleagues. In a series of studies, they demonstrate that the decline in the ability to acoustically discriminate non-native contrasts occurs within the first year of life (Werker, Gilbert, Humphrey & Tees, 1981; Werker & LaLonde, 1988; Werker & Tees, 1984 a, b; and also Best & McRoberts, 1989; Best, 1994). What is particularly fascinating is that this decline in perceptual capacity does not appear to be temporally uniform for all non-native contrasts. Experimental results indicate that perceptual sensitivity to certain non-native contrasts is lost before sensitivity to others. In a longitudinal study, using a head-turn paradigm, Werker & Tees (1984a) showed that six to eight month old English infants were able to perceive the non-native contrasts between alveolar and retroflex stops /t/ and /ʈ/ (Hindi) and velar and uvular stops /k/ and /q/ (Interior Salish, Nthlakampx). When these infants were slightly older (eight to ten months), they were still able to perceive the Hindi contrast but had lost the ability to perceive the Salish contrast. At the age of ten to twelve months old, all of the English infants were unable to perceive either non-native contrast, while Hindi and Salish infants of the same age were still able to discriminate their respective native contrasts.[6] What these studies suggest is that loss of perceptual sensitivity to non-native contrasts is gradual and proceeds in a systematic order.

Several attempts have been made to explain this decline in perceptual sensitivity. It has been suggested that if specific phonetic experience is not forthcoming within some critical period (which appears to be the first year), then the phonetic feature detectors that subserve the contrast in question undergo a loss of sensitivity (Eimas,

---

[6] This developmental sequence was first established by a cross-sectional study using English infants and the Hindi and Salish contrasts. However, in the cross-sectional study, one group of infants was tested on the Hindi contrast and a different group was tested on the Salish contrast, so we can not directly compare performance on the two contrasts to determine what the relative order of perceptual decline within a given child was.

1975; Tees & Werker, 1984). An English child, for example, never hears a contrast between alveolar and retroflex stops and, therefore, loses the ability to perceive that contrast. Under this account, perceptual sensitivity is lost as a direct result of the nature of the input itself (i.e., lack of a contrast in the input), rather than some perceptual re-organization taking place in the learner. One prediction of this position is that perceptual sensitivity will not be lost for segments that vary allophonically in the native language since the infant will have phonetic experience with both segments. The research reported below clearly refutes this prediction (at this point, it is not clear whether there is a different perceptual effect for allophones that are in complementary distribution and those that are in free variation). Another prediction of this position is that perceptual sensitivity will be lost for every contrast not manifested in the input. However, work by Best, McRoberts & Sithole (1988) demonstrates that perceptual sensitivity is not lost for every non-native contrast. These researchers found that English infants (ranging in age from six to fourteen months) as well as English adults accurately discriminate three sets of Zulu clicks in an auditory task.[7] What is unique about this particular non-native contrast is that neither member of the Zulu contrast occurs in the English inventory. Recall that in Werker's studies, one member of each contrast investigated is an English phoneme (alveolar stop /t/ and velar stop /k/). Thus, it appears that perceptual sensitivity to a non-native contrast is only lost if one of those segments occurs in the learner's native language; sensitivity is retained if neither segment occurs in the L1.[8]

To explain this phenomenon, both Werker and Best have tentatively suggested that the decline in ability to discriminate some non-native contrasts may reflect the first stage in the phonological development of the child. Although neither author is specific as to which aspect of the developing phonology might be responsible for this loss, their suggestion raises an intriguing possibility that I will explore here.

---

[7] Clicks are ingressive velaric consonants (unlike any English consonant) produced by the formation of a suction chamber in the oral cavity followed by an abrupt release of negative pressure at the blade, tip or side of the tongue, or at the lips (Catford & Ladefoged, 1968).

[8] Later, I will revise this position and propose that sensitivity to a non-native contrast is retained if neither segment occurs in the L1 OR if the *feature* that underlies the non-native contrast is employed in the L1.

3.2.2. The Role of the L1 Phonology in Speech Perception: A Proposal

When we consider the findings of Werker & Tees and Brown & Matthews together, an interesting parallelism emerges: infants perceptual capacities gradually *"degrade"* from all possible contrasts to only native contrasts (with the exception of certain non-native contrasts, to which I will return) while their ability to discriminate segments phonologically gradually *improves* from no contrasts to only native contrasts. Thus, there is a convergence of both types of capacities on the learner's set of native sounds. Moreover, when we examine the specific orders of decline and improvement, we find that they mirror one another. Although such a comparison is limited at this point, due to the few number of non-native contrasts that have been investigated thus far, it is worth pursuing. According to Brown & Matthews, children first phonologically differentiate labials from velars, followed by labials from coronals. They do not distinguish segments that require a Coronal node, such as /l/ and /r/, until relatively late.[9] So, the node that distinguishes velars segments from other places of articulation is posited by the child before the node to distinguish coronal segments is posited. Measuring auditory perception, Werker found the reverse order, with the perception of contrasts involving velars declining before contrasts involving coronals.[10] It is plausible, then, that the very construction of a segmental representation is responsible for the degradation of acoustic speech perception.

I propose that the elaboration of feature geometry in the child's grammar imposes upon his or her perception the specific boundaries within which phonemic

---

[9] Although English contains other coronal sounds (e.g. /t/ or /n/), it is general assumed that in the absence of a phonemic contrast within the class of coronal sounds, coronals will be represented in the phonology with a bare Place node (see papers in Paradis & Prunet, 1991; cf. McCarthy & Taub, 1992). The coronal feature will, then, either be inserted during the phonetic component of the grammar or the bare Place node will be interpreted directly with a coronal place of articulation.

[10] One curious fact is that there appears to be a time lag (approximately four to six months) between the age of the perceptual loss and the corresponding phonological development. A possible explanation for this time lag is that there is a confound between lexical development and phonological development, such that segmental representations are integrated into lexical items (which is what Brown & Matthews actually measured) shortly after they are first acquired. If this is indeed the case, acoustic discrimination tasks might provide a means of measuring the phonological development of children at even earlier ages than is currently available.

categories are defined.[11] In other words, the degradation of the perceptual capacities and the increase in the ability to distinguish sounds phonologically are the result of the same internal mechanism, namely the construction of phonological representations. This layer of phonological structure then mediates between the acoustic signal and the auditory processing system. The phonological structure can, thus, be viewed as a filter which funnels acoustically distinct stimuli into a single phonemic category. Best's investigation of the perception of Zulu clicks by English infants and adults provides support for this proposal. Recall that, under the "feature detector" explanation, it is surprising that English speakers are able to perceive the differences between the various clicks, since they have had no exposure to these sounds.[12] If, on the other hand, perception declines as a result of the elaboration of phonological structure, this is exactly what we would expect. English infants never hear these types of segments used contrastively in the input, so no phonological structure has been posited. Since no phonological representation intervenes between the acoustic signal for these segments and the auditory processor, clicks will be discriminated purely on the basis of their acoustic characteristics.

The schematized diagram in Figure 1 illustrates the role of the intervening layer of phonological structure and how this level, in effect, funnels the acoustic signal into the phonemic categories of the speaker's language.[13]

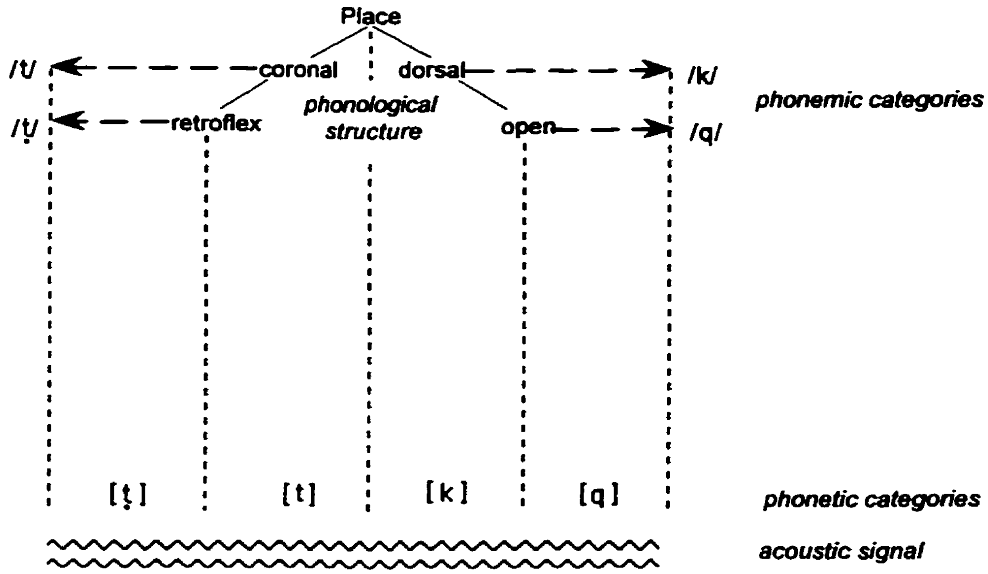
---

[11] A slightly weaker position is that, although the degradation of speech perception is not *caused* by the acquisition of phonological representations, it is, in fact, that phonological structure which ultimately enforces the boundaries of categorical perception.

[12] Although English learners may have experience with clicks as non-speech sounds (e.g. tsking or clucking to urge a horse along), this does not constitute linguistic experience.

[13] Best (1993, 1994) has developed a Perceptual Assimilation Model to explain the role that a speaker's L1 phonological system plays in the perception of non-native sounds. According to her model, non-native sounds are assimilated to a listener's native categories on the basis of their respective articulatory similarities; the degree to which a non-native contrast can be assimilated to native categories determines how well (if at all) a listener will be able to perceive that non-native contrast. While Best's proposal is similar in spirit to the one outlined in this paper, it lacks precise objective criteria for determining (and predicting) how non-native contrasts will be assimilated into native categories. For example, both Best's model and my proposal predict that English speakers will be unable to discriminate the Salish /k-q/ contrast, however according to Best this inability results from the fact that the two sounds are "articulatorily similar" for English speakers, whereas under my proposal, the two sounds are not discriminated because English speakers' grammars lack the phonological feature that distinguishes the two sounds (either, [high] or [pharyngeal]). Thus, although Best's model describes the role of a speakers L1 phonological system plays in the perception of non-native sounds, it does not provide an explanation.

Figure 1
Mediation of Speech Perception by Phonological Structure



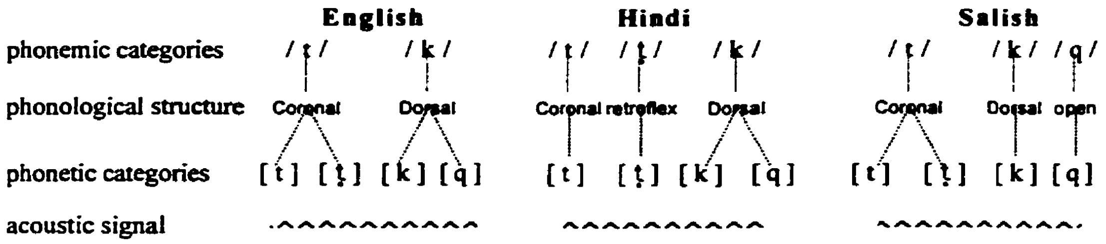
The feature geometry depicted is from a hypothetical language in which /t/, /ṭ/, /q/ and /k/ are distinct phonemes; we will only consider the Coronal and Dorsal nodes (and their dependents) of the geometry.[14] Starting from the bottom of the diagram, the acoustic signal is first broken down into phonetic categories. At this level, the acoustic signals for an alveolar [t] and a retroflex [ṭ] remain distinct. These stimuli then pass to the second level which consists of a speaker's feature geometry. This phonological structure serves to further categorize the phonetic stimuli into phonemic categories. Because this language exploits a dependent feature of [Coronal], the phonetic signals for [t] and [ṭ] are channeled into the distinct phonemic categories /t/ and /ṭ/, which are then fed into the auditory processor. The acoustic signals for [q] and [k] are processed in the same way. This model of speech perception is supported by research by Werker & Logan (1985), who found evidence for three distinct levels of processing: depending on the length of the interval placed between the stimuli (and hence on the memory load required to perform the task), subjects exhibit perception at either the auditory, phonetic or phonemic level. In particular, these researchers showed that, under certain

[14] For this illustration and subsequent argument, it is not important whether [high] or some other phonological feature (e.g. [pharyngeal] distinguishes /q/ and /k/).

conditions, English speakers are able to auditorily discriminate the Hindi /t–ṭ/ contrast more accurately than predicted by chance. According to the model in Figure 1, the acoustic signal will first be divided into distinct phonetic categories, which are only subsequently categorized into native phonemic categories. Thus, regardless of the phonological system of a speaker, non-native contrasts are distinct at some level and may be discriminated under certain controlled conditions, as Werker & Logan have demonstrated.

The diagram in Figure 2 illustrates, in slightly more abstract terms, how the speech perception of English, Hindi and Salish speakers differs from one another.

Figure 2
Cross-language Speech Perception



In English, the feature [Coronal] serves to distinguish coronals from non-coronals (e.g., /t/ vs. /p/), but no distinction is made within the coronal place of articulation (e.g., /t/ vs. /ṭ/). Thus, the English feature geometry does not contain the feature [retroflex]. As a result, different coronal sounds, regardless of their distinct acoustic signals, are perceived as a single phonemic category. Likewise, English makes no phonemic distinction between velar and uvular sounds; therefore, the Dorsal node has no dependents and velar and uvular sounds will be perceived as the English phoneme /k/. The feature geometry of Hindi also lacks the Dorsal dependent [open], so that perception of [k] and [q] as /k/ is the same as for English speakers. Unlike English, however, the Hindi feature geometry contains both the feature [Coronal] and its dependent [retroflex]. Thus, all coronal sounds will not be funneled into one phonemic category; the two features in the geometry ensure that /t/ and /ṭ/ will be perceived as distinct phonemes. In Salish, the situation is just the reverse: /k/ and /q/ are perceived as distinct phonemes (due to the presence of the feature [open] as a dependent of the Dorsal node) while the acoustic signals for /t/ and /t̪/ are perceived as the single phoneme /t/ (due to the absence of [Coronal] dependents).

To summarize thus far, I have suggested that a learner's acquisition of feature geometry in L1 acquisition causes the gradual decline in the ability to acoustically discriminate non-native contrasts. Exposing the link between a learner's phonological development and his or her subsequent perception of speech will have important implications for the acquisition of a second language. In particular, we have established how the L1 grammar maps the L2 input onto existing L1 phonological categories, effectively eliminating cues in the acoustic signal that could potentially trigger further acquisition. In the next section, I outline some of the predictions this model of phonological interference makes for the L2 acquisition of non-native contrasts.

3.2.3. Implications for L2 acquisition

There are three primary types of non-native contrasts that an L2 learner might acquire. In the first type, each member of the contrast is similar to distinct segments in his or her L1. For example, the Salish contrast between glottalized /t'/ and glottalized /k'/ corresponds to the English /t/ – /k/ contrast. In such a case, the learner will categorize the members of the non-native contrast into phonemic categories of his or her L1, which are themselves contrastive. Thus, an L2 learner will be able to acoustically discriminate a non-native contrast if the members of that contrast correspond to distinct *phonemes* in the learner's native language. An L2 learner in this situation is able to acoustically discriminate the non-native contrast, not on the basis of their individual phonetic characteristics, but by virtue of the fact that each is perceived as a distinct phoneme in the L1. Therefore, no new phonological structure will be added to the interlanguage grammar and the speaker will likely utilize his or her existing representations of the native segments to represent the new segments.[15] Acquisition of

---

[15] In order to determine whether the learner has constructed new representations or has merely substituted existing ones, one would need to examine whether the speaker could distinguish

this type of non-native contrast is not particularly illuminating to our understanding of how (if at all) novel segmental representations are acquired by L2 learners, since the auditory and phonological discrimination of this type of contrast can be accomplished using one's L1.

A more interesting situation occurs when neither member of the non-native contrast is (or corresponds to) a phoneme in the learner's native language. The acquisition of Zulu clicks by English speakers falls into this category. No phonological structure will have been posited in the learner's L1 to represent either of the segments, so that perception of the contrast will not be blocked by the learner's L1. If UG is still available for L2 acquisition, then detection of the contrast will trigger the addition of structure to the learner's feature geometry. He or she will, therefore, be able to construct the representations of the two non-native segments. If UG is not available, then, although the learner will detect a contrast in the input, the mechanism for elaborating segmental structure will not be operative and novel segmental representations will not be constructed. Acquisition of this type of contrast is the most interesting, as it most closely mirrors first language acquisition. Unfortunately, given the linguistic constraints on possible phonological oppositions, it is very difficult to find a phonemic contrast that does not share either member with some other language (clicks are a rare example).

The final type of non-native contrast that a L2 learner might try to acquire is the type studied by Werker, in which one of the members of the non-native contrast is a phoneme in the learner's L1. In this case, the proposal outlined above predicts that perception of the non-native contrast will be blocked by the learner's L1 grammar: the L1 feature geometry will funnel the acoustic signals for the two segments into one native phonemic category. If the learner does not (in fact, *cannot*) detect a contrast between two segments in the L2 input, then elaboration of the feature geometry in the learner's interlanguage grammar will not be triggered. Hence, the learner will not have the phonological structure necessary to represent the two segments as distinct phonemes

---

(perceptually and linguistically) the novel segments from the corresponding native segment (e.g. /t'/ from /t/).

in his or her interlanguage grammar. The remainder of this paper reports two experimental studies which investigate the predictions concerning acquisition of this type of non-native contrast.

3.3 Experiment 1

3.3.1. Properties of Japanese and Mandarin Chinese & the contrasts investigated

This experiment investigates the acquisition of the English /l–r/ contrast by Japanese and Mandarin Chinese speakers. These language groups were chosen because neither contrasts /l/ and /r/ phonemically. In addition, an extensive amount of research has already been conducted on the perception of /l/ and /r/ by Japanese speakers (e.g., Goto, 1971; Miyawaki, Strange, Verbrugge, Liberman, Jenkins & Fujimura, 1975; Sheldon & Strange, 1982; Yamada, 1995; *inter alia*). This previous research demonstrates clearly that Japanese speakers do not accurately perceive the /l–r/ contrast; thus, we have a strong indication that our L2 learners have difficulty with this English contrast. However, the prior research has only assessed their auditory perception of these sounds and has not investigated whether they can acquire novel phonological representations for these sounds. Hence, research that assesses both the auditory and phonological capabilities of Japanese speakers is still needed; to the best of my knowledge, the perception and acquisition of the /l–r/ contrast by Chinese speakers has not yet been investigated.

Before examining the design and results of this experiment, we must first consider the characteristics of the language groups being investigated in closer detail. The phoneme inventory of Japanese contains one liquid described as a flap /ɾ/, which is not identical to the central approximant /r/ in English (Maddieson, 1984). This phoneme has several variants, conditioned by the phonological context. One of these variants is phonetically similar to English /l/. Thus, Japanese phonetic /l/ and /ɾ/ are allophones of a single liquid phoneme /ɾ/. The underlying representation of this phoneme is given below, repeated from (3).

(4) /ɾ/

ROOT

SV PLACE

|

approximant

Since the two liquids are not separate phonemes in Japanese, /ɾ/ need not be distinguished from /l/ underlyingly. Thus, in contrast to English /r/, the feature [Coronal] is not present in its representation (cf. example (2b). This Japanese liquid may only appear in a simple onset position; consonantal clusters are prohibited by Japanese syllable structure constraints and although codas are permitted, they may only be filled by the coronal nasal /n/, a nasal that is homorganic with the following onset or the first half of a voiceless obstruent geminate. These phonotactic constraints further distinguish Japanese from English and may, in fact, factor into the acquisition of English phonemes.

Mandarin Chinese, like Japanese, lacks the contrast under investigation; however, its inventory contains the lateral approximant /l/ rather than the central approximant.[16] The representation of this segment is identical to the Japanese liquid shown in (4). The phonetic realization of this segment does not vary between /l/ and /r/, as it does in Japanese. Consonantal clusters are also prohibited in Mandarin Chinese and codas are restricted to the coronal and velar nasals. Thus, Mandarin Chinese and Japanese are quite similar with respect to their inventory of liquids and syllable structure constraints.

---

[16] The claim that Mandarin Chinese does not contrast /l/ and /r/ phonemically requires some comment. This language contains /l/ and a segment which is transcribed in romanized script as "r"; this transcription gives the impression that there is a contrast between the lateral approximant /l/ and a central approximant /r/. This "r" segment, however, is classified by linguists as a voiced retroflex fricative, /ʐ/. For this study, I follow Maddieson (1984) in treating /ʐ/ as a voiced retroflex fricative and, crucially, not as a retroflex sonorant. But compare Rice (1992) who analyzes this segment as /r/ underlyingly; postulating that it surfaces as a voiceless retroflex fricative [ʐ] in onset position and as rhoticization of the vowel when in coda position [ɚ]. Thus, according to this analysis, /r/ and /l/ do contrast as sonorants in Mandarin. However, it is not clear that this analysis is correct. The coda position in Mandarin is restricted to nasals; thus it is unlikely that the rhoticization of the vowel is from the presence of an approximant in the coda position. Finally, only certain vowels are rhoticized (Chao, 1968; Wu, 1991). This suggests that the rhoticization is a property of the vowel itself, rather than the result of /r/ in the coda position.

According to our proposal that a speaker's L1 feature geometry causes the decline in his or her ability to perceive non-native contrasts, the Mandarin Chinese and Japanese speakers should both be unable to acoustically discriminate /l/ from /r/. Given that the perception of a contrast is required for the acquisition of segmental structure to represent that contrast, speakers of both languages should also be unable to acquire this non-native contrast.

3.3.2. Methodology

Subjects

Thirty–two adults with no evidence of hearing loss participated in this study. Two of these subjects were subsequently eliminated as outliers based on their language profiles; their performance is discussed in footnote 26. The first experimental group consisted of ten Japanese speakers who had learned English as their only second language. All of the subjects in this group were raised in Japan and came to North America as adults to study in undergraduate or graduate programs at McGill University in Montreal, Canada. These subjects were between twenty-five and thirty-seven years of age at the time of testing. The reported age of first exposure to English for these subjects ranged from ten years old to sixteen years old. Each subject had studied English in school in Japan for a minimum of six years up to a maximum of thirteen years. The length of time these subjects had been continuously living in North America ranged from two months to seven years.

The second experimental group consisted of ten Chinese speakers who had learned English as a second language (nine of these were native Mandarin speakers and one was a native speaker of Cantonese who learned Mandarin in grade school). Chinese and English were their only languages, with the exception of one Mandarin speaker who also spoke Japanese as a second language. All of the subjects in this group were raised in either China, Taiwan or Hong Kong and came as adults to North America to study or work at McGill University. They were between twenty-five years and fifty-three years old at the time of testing. The reported age of first exposure to English varied considerably among the Chinese subjects, from as young as four years old up to thirty-eight years old.[17] The number of years these subjects had spent studying English in China also varied considerably, ranging from one and a half years to thirteen years. These subjects had lived in North America from eight months to nine years at the time of testing.

The final group in this study consisted of ten monolingual speakers of American English who served as controls. They ranged in age from twenty-one years to fifty-three years. Table 1 summarizes the background information for each of the three groups. Note that the Japanese and Chinese groups are quite comparable in terms of age and years spent in North America, while the Japanese subjects have spent more time studying English in their country of origin and were exposed to English at an earlier age than the Chinese subjects. All subjects were paid for their participation.

Table 1
Subject Information

Group	Mean Age at Testing	Mean Age of Exposure	Mean Years Studied	Mean Years in N. America
Japanese	30.3	12.3	9.6	3
Chinese	32.4	16.7	5.6	4.5
Controls	31.5	—	—	—

Tasks & Materials

Rationale for tasks

There are two types of data that might be used to determine whether a second language learner has acquired a non-native contrast. Production data might provide evidence that the learner has acquired sufficient underlying structure to represent a given contrast if the learner accurately and consistently produces that contrast. However, as Brown &

[17] The Chinese subject that was exposed to English at four years old does not technically qualify as an adult L2 learner. However, as reported in §4.3.1 and §4.3.2, there was no correlation between age of exposure and performance on the two tasks in this study. Moreover, this young subject did not perform significantly differently than the other Chinese subjects and was, therefore, included in my analysis.

Matthews (1993, in press) point out, using production data alone to evaluate the phonological development of *children* is potentially misleading because a child's production of a particular contrast may be delayed due to independent limitations (e.g., insufficient motor control to make the necessary articulatory gesture). A child may have the necessary structure in his or her grammar to distinguish two sounds yet be unable to produce the difference between the two sounds consistently or even unable to produce the sounds at all, in the case of pre-verbal children. Thus, relying on production data may lead to an underestimation of the child's phonological competence.

In the case of second language acquisition, the situation may be just the reverse. Several researchers have demonstrated that L2 learners may be able to accurately produce a non-native contrast even though the same learners are unable to distinguish the two sounds perceptually (Brière, 1966; Flege, 1995; Goto, 1971; Sheldon & Strange, 1982). Since adult learners have a developed motor control system, they are often able (with practice) to execute the necessary articulations. In fact, in the case of the Japanese speakers, much of their phonological training in English classes is devoted to the mastery of the *articulation* of /l/ and /r/. Once a speaker knows the spelling of a word that contains /l/ or /r/, he or she can accurately produce the correct liquid, thus giving the appearance of having acquired the contrast.[18] Therefore, if we rely on production data we may falsely attribute more segmental structure to a learner's underlying phonological competence than he or she actually has. On the other hand, some L2 learners are unable to correctly produce a novel contrast in spite of their ability to perceive that contrast, in which case we would underestimate the learner's competence.

Comprehension data also provide an opportunity to determine whether a learner has developed sufficient underlying structure to support a phonemic contrast. Data from comprehension tasks are taken by many researchers to be a more accurate indication of the underlying competence than production abilities, precisely because fewer peripheral

---

[18] To my knowledge, no one has investigated how this knowledge of proper articulation might be encoded into the learner's lexical representation of words. It is not clear that this knowledge, which is dependent on orthography, is linguistic in nature or is represented in terms of phonological structure.

performance systems intervene which might obscure the data. If a learner accurately and consistently distinguishes two particular sounds in a comprehension task, then we can assume that the learner has the structure in his or her grammar necessary to represent the particular contrast. On the other hand, if a learner does not distinguish two sounds in this type of task, we can assume that his or her phonological representation lacks the relevant structure to differentiate the two segments.

AX Discrimination

There are two aspects of the L2 learners' interlanguage grammar that will be investigated here. Since our hypothesis is that the learner will only acquire the segmental representations for those contrasts which he or she accurately perceives, we need to first determine whether or not the learner can acoustically discriminate the two sounds that comprise the non-native contrast. An AX Discrimination task was used to assess the subjects' ability to acoustically discriminate /l/ from /r/. In this task, the subject hears a minimal pair, one item containing an /l/ and the other item containing an /r/, and is asked to indicate whether the words are the same or different (e.g., *rip / lip*). The items used in the test were natural tokens of real English monosyllabic words. These tokens were spoken by a man with a standard (Californian) American English accent and recorded so that the stimuli were identical for every subject. Three different types of words were used for the experimental conditions: the first type differed only in a single onset consonant (e.g., *rock / lock*), the second type differed in complex onset (e.g., *crown / clown*) and the third type differed with respect to the coda consonant (e.g., *ear / eel*). The position of the /l/ and /r/ was varied in order to determine whether the position of the contrast in the word would be a factor in the ability of the learners to acquire the contrast. In particular, segments in the coda position of a syllable have a unique relationship with the nucleus since they are both in the rhyme of the syllable and can, therefore, affect the phonetic realization of the nucleus; for example, it has been established that voiced consonants systematically lengthen the duration of the preceding vowel (Peterson & Lehiste, 1967). Therefore, we might expect differential performance on the coda condition, quite independent of the phonological properties of the learner's L1.

Two types of foils were included in the test materials. Seven minimal pairs contained onsets that differed by a consonant other than /l/ or /r/ (e.g., *peas/keys*). This foil type was included as a means of checking that poor performance on the task was not due to difficulty with the task itself; any difficulty with the task would be reflected in poor performance on the foil items as well as on the experimental items. These foils were also intended to distract the subjects' attention from the specific contrast on which they were being tested. A second type of foil was included to detect any response bias. Since each experimental minimal pair differs with respect to some consonant, the correct response for every trial is that the words are acoustically different. Any response bias or strategy toward responding that all of the pairs were different would result in accurate performance. This was not an unlikely possibility, especially for the Japanese speakers who are (explicitly) aware that /l/ and /r/ are contrastive in English. If a subject had guessed what the experiment was investigating, he or she might have responded "different" to each of the test items. This second foil type consisted of sixteen pairs of words in which the two stimuli were identical (e.g., *lamb/lamb*, *glass/glass*, *pear/pear*). If a subject responded that these identical stimuli were different a majority of the time, that subject's data were discarded.

Six pairs of each of the three experimental conditions (e.g., simple onset, complex onset, coda) were included. In order to examine the effect (if any) of the neighboring vowel on the acquisition of the contrast, the vowel was varied in each of the pairs, when possible.[19] The initial consonant of the complex onset also differed

---

[19] An anonymous reviewer has pointed out that three of the six coda pairs are not perfect minimal pairs, since the vowels in these pairs are not identical. As mentioned above, a coda /l/ or /r/ will have a significant effect on the phonetic realization of the preceding vowel; thus, regardless of which vowel is in the nucleus, the coda liquid will color its realization. The degree to which the liquid colors the nucleus differs considerably depending on the type of vowel; so, for example, back vowels such as /o/ will be more strongly affected than front vowels such as /i/ (compare bowl/boar with tile/tire). As the list of stimuli in Appendix A indicates, both back and front vowels were both included to ensure that the obtained results were not simply an artifact of a single vowel type; unfortunately, one of the consequences of controlling for vowel height and backness is that some of the minimal pairs are not perfect. However, as we will see in the results section, subjects performed uniformly on all of the coda minimal pairs, indicating that our findings are not compromised by the inclusion of the minimal pairs

with respect to place of articulation and manner of articulation whenever possible. These variations ensured that any difficulty with the contrast was not due to difficulty with a specific vowel or type of consonant in the complex onset. The vowels were varied for the foil items as well. Care was also taken to choose items that would be relatively familiar to second language learners. See Appendix C for a complete list of stimuli used. Each of the experimental and foil pairs appeared twice in the test for a total of sixty-six items. The order of the two stimuli in a given trial was reversed for each of the two trials containing that pair in order to guard against any possible effects for presentation order of the contrast (e.g., once as *rake/lake*, once as *lake/rake*). The order of the trials in the test was pseudo-random. If any sequence of three identical conditions (i.e., word-positions) occurred after randomly ordering the items, the third item was separated from the other two. This order of the sixty-six trials was then reversed to create a second test version to control for potential fatigue factors. Each test version was presented to fifteen of the thirty subjects.

Forced Choice Picture Selection

The second aspect investigated is whether those subjects who can acoustically discriminate the non-native contrast are able to acquire the phonological structure necessary to distinguish the two sounds phonologically; in other words, when the requisite conditions are met, is acquisition of non-native phonemes possible? Subjects were given a Forced Choice Picture Selection task (modified from Brown & Matthews, 1993, in press), in which the subject is presented with two pictures as well as a verbal cue that corresponds to one of the pictures.[20] For example, the subject would see a picture of a *rake* on the left side of the page and a picture of a *lake* on the right side and, at the same time, hear the word *lake*. The subject's task is to indicate which of the pictures the verbal cue names. In order to successfully complete this task, the learner must access his or her internal phonological representations of the pictured objects and

---

whose vowels are more strongly affected by the coda. Phonetic transcriptions of the coda stimuli, as produced for the experimental task, are included in Appendix A.

[20] The materials used in stimuli preparation for the picture test reported here were adapted from the Bilingual Aphasia Test (Paradis & Libben, 1987).

determine which lexical representation corresponds to the verbal stimulus. If the subject's lexical representations of the pictured objects are identical (i.e., if they do not have the necessary phonological structure to contrast /l/ and /r/), then he or she will be unable to determine which picture the verbal cue corresponds to and should perform the task with chance accuracy. Successful completion of this task indicates that the subject has acquired the non-native contrast and differentiates /l/ and /r/ in his or her lexical representations.

Each stimulus item contained two black and white pictures (see Appendix D). The pictured objects differed minimally with respect to the /l/ - /r/ contrast. Each array of pictures was presented with a verbal cue that corresponded to one of the two pictures. The monosyllabic words used in this task were the same as those used in the AX Discrimination task, with the exception of the identical-word pairs, which were omitted. As in the discrimination task, there were six pairs of each experimental condition and seven foil pairs which contrasted segments other than /l/ and /r/. Each of these pairs was repeated in the test for a total of fifty trials. Each occurrence of a particular diad was presented with a different verbal cue and each time in a different spatial arrangement. In order to avoid any biases for position on the card, twenty-five trials had the correct picture in the right position and twenty-five had the correct picture in the left position; thus, in half of the trials the pictured object in the right position contained an /l/ and in the other half an /r/. The order of the trials in this test was, as in the discrimination task, pseudo-random. If any sequence of three identical conditions occurred after randomly ordering the items, the third item was separated from the other two. Trials for which the correct answer was consecutively three or more times in the right position or the left position were also separated. This order of the fifty picture arrays was then reversed for a second test version, which was completed by half of the subjects.

A training book was constructed which included every picture (one to a page) that appears in the experimental test. This book was used to familiarize the subjects with each of the pictures, and corresponding name, to be shown in the Picture task. This was done in order to minimize any errors that might be caused by the subjects' unfamiliarity with a particular stimulus or illustration of an item.

Procedure

Each subject was tested individually in a quiet room with the experimenter. Both tasks were completed in one testing session which lasted approximately forty minutes, including a five minute break between tasks. After filling out a brief language profile, each subject completed the Discrimination Task followed by the Picture Task. The discrimination task was given before the picture task to ensure that the subjects discriminated the minimal pairs on the basis of their acoustic properties alone, rather than any knowledge of the particular items (i.e., the knowledge that a word containing /l/ had a minimal pair containing /r/, or vice versa). The instructions for each task were read to the subject in English and appeared in written form on each response sheet. An example was also given before testing began. The stimuli were presented to the subjects binaurally over a Sony business machine (BM-75) at a comfortable listening level. Subjects responded immediately to each trial and were not allowed to rewind or stop the tape. Subjects responded to each trial by circling either "same" or "different" on the response sheet or, in the event that they could not make a judgement, "not sure". This task took approximately ten minutes for subjects to complete.

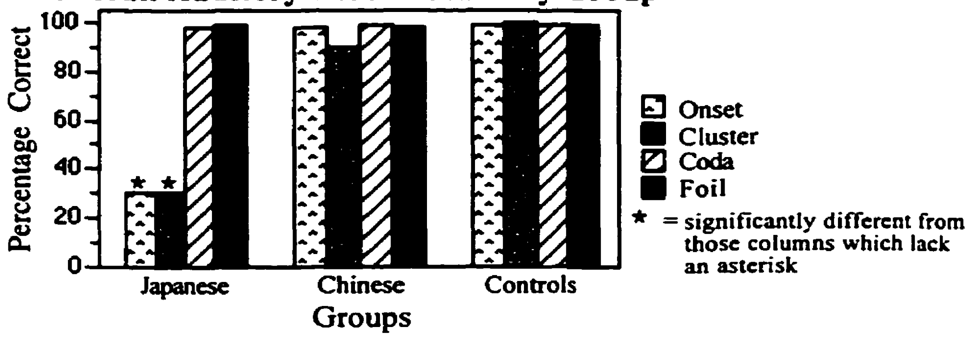
Before starting the Picture task there was a short training period during which the experimenter showed each of the pictures in the training book to the subject and told him or her the name of each object. If the subject did not know the name or did not recognize the object, a few moments were spent explaining the meaning of the stimulus item to the subject. When the experimenter was satisfied that the subject was familiar with each of the pictures, the Picture task was administered. Subjects were instructed to circle either "left" or "right" on the response sheet, depending on the position of the named object, or "not sure" in the event that they could not determine which picture corresponded to the verbal cue. This task took approximately fifteen minutes to complete. At the end of the testing session, the subject was asked informally if they had found either of the tasks difficult and whether or not they had experienced any difficulty learning the /l/–/r/ contrast in English.

3.3.3. *Results*

Auditory Discrimination Task

For the auditory discrimination task, a response that the two words in the minimal pair were "different" was counted as correct and a "same" response was counted as an error. Performance on foil pairs which consisted of identical words (e.g., *rice/rice*) was not included in the tabulation of subjects' scores since accurate performance on these items would be consistent with the subject's L1 and would not reveal properties of his or her interlanguage grammar. "Not sure" responses on this task were discarded and a subject's overall performance score was calculated as the percentage correct on the remaining experimental trials.[21] Performance scores on each of the experimental conditions (i.e., different positions of the /l/–/r/ contrast in the word) were tabulated separately for statistical analysis. No effect was found for test version, so scores from versions A and B were pooled. Figure 3 reports the mean performance scores of all groups, according to experimental condition.

Figure 3
Overall Auditory Performance By Group



[21] A subject may have responded "not sure" for a variety of extralinguistic reasons, including having not heard one of the two items completely. Importantly, a "not sure" response would not reflect either the subject's L1 or interlanguage grammar: as both grammars would categorize the two acoustic signals as either the same sound or as different sounds, there should not be confusion or uncertainty.

From the subjects' near perfect to perfect performance on the foils (i.e., non /l/-/r/ pairs), it is clear that poor performance on any other experimental condition is due entirely to the /l/-/r/ contrast and not the task itself. A factorial ANOVA reveals highly significant differences between groups for the onset condition, $F(2, 27) = 171.025$, $p = .0001$, and the cluster condition, $F(2, 27) = 71.381$, $p = .0001$. Post-hoc Scheffe tests ($p<.05$) indicate that the Japanese group differs significantly from both the Chinese and Control groups on these two conditions while the Chinese and Control groups do not differ significantly from each other. There was no significant difference between the three groups on the coda condition or the foils. In order to evaluate the performance of each language group, further analyses were carried out on individual groups.

Japanese

Performance scores for the Japanese group are given in Table 2; these scores are reported in percentage correct. A repeated measures ANOVA on the mean scores shows a significant difference between the various experimental conditions, $F(9, 30) = 185.017$, $p = .0001$. Post-hoc Scheffe tests ($p<.05$) indicate that performance on the coda condition is significantly better than on either the onset or the cluster condition whereas performance on the onset and cluster conditions are not significantly different from one another. These results suggest that Japanese learners of English cannot perceive the contrast between /l/ and /r/ auditorily when it is in the onset (31%) or cluster position (38%).[22] What is somewhat surprising is the Japanese subjects' performance on the coda position which, at 99% correct, is strikingly different from the two other experimental conditions.

---

[22] Based on the Japanese L1 grammar, we would theoretically expect 0% correct performance on these conditions. That is, they should not discriminate any of the stimuli pairs. However, in practice, it is possible that some of the stimulus items were discriminated on the basis of some non-linguistic criterion, such as duration or amplitude (the stimuli were not controlled for these types of variations). In any case, the below-chance performance indicates that the Japanese subjects were not able to reliably use this non-linguistic criterion to distinguish the stimulus pairs. Moreover, their performance demonstrates that they are unable to utilize the phonetic cues (which are consistently present) to distinguish /l/ and /r/.

Table 2
Japanese Performance Auditory Task

Subject	Onset	Cluster	Coda
1	25	33.3	**100**
2	20	30	**100**
3	40	40	**100**
4	58.3	41.7	**100**
5	33.3	58.3	**100**
6	9.1	11.1	**100**
7	50	41.7	**100**
8	16.7	50	**100**
9	25	41.7	**92.9**
10	33.3	33.3	**100**
Means	31.1%	38.1%	**99.3%**

While group means indicate a difference between Chinese and Japanese speakers, we must consider the performance of individuals to ensure that differences between the groups are not due to a few individuals. Based on a standard binomial distribution, the probability that an individual could randomly choose correctly between two equally likely choices ten times out of twelve is less than the standardly accepted 5%. Therefore, a score of 83.3% (i.e., ten out of the twelve trials) has been set as the criterion for successful performance on a given experimental condition. Bolded typeface in Table 2 (and all following tables) designates those conditions on which criterion was reached. Each Japanese subject reached criterion in the coda condition, yet failed to reach criterion on both the onset and cluster conditions. This indicates that, although there is some variation (compare subjects 4 and 6), the group scores do, in fact, reflect individual performances.

Chinese

Table 3 shows the performance of the Chinese subjects. A repeated measures ANOVA on the mean scores reveals a statistical difference between experimental conditions, $F(9, 30) = 3.332$, $p = .0343$. Post-hoc procedures indicate that performance on the cluster condition is significantly different than performance on the onset and coda conditions according to the Fisher test (however, not according to the more stringent Scheffe procedure). What is important to note about these Chinese results is that performance on all of the experimental conditions is well above the level predicted by chance. This is in stark contrast to the performance of the Japanese subjects on the onset and cluster conditions. Thus, although performance by the Chinese is significantly better on the onset and coda conditions, performance on the cluster condition is still well above chance.

Table 3
Chinese Performance Auditory Task

Subject	Onset	Cluster	Coda
11	91.7	50	91.7
12	100	100	100
13	100	91.7	100
14	100	100	100
15	100	91.7	100
16	83.3	66.7	91.7
17	100	91.7	100
18	100	100	100
19	100	100	100
20	100	100	100
Means	97.5%	89.2%	98.3%

Examination of the individual scores shows that every subject, except for two, reached criterion on all three conditions. Subjects 11 and 16 failed to reach criterion on the cluster condition. Apart from these two exceptions, the individual data are consistent with the group results: each Chinese subject is able to perceive the /l/–/r/ contrast auditorily in onset position and coda position, while the majority of subjects are also able to perceive that contrast in consonantal clusters.

These results show that, contra to our hypothesis, the two language groups perform differently on this task: Chinese speakers accurately discriminate the contrast while Japanese speakers perform at or below chance. This difference in language groups suggests that the subjects' specific L1 (not simply whether their L1 contains the contrast under investigation) is the key factor in determining performance on the task. We would also like to determine if any other factors play a role in the performance of this task. In order to investigate these factors, several correlations were computed across groups and within groups. The first factor investigated was age of exposure. We found a slight correlation between age of exposure and performance on the auditory task; however, this correlation was not significant, $r = .285$, $p > .05$. Similarly, when we factor out the L1s, and consider each group individually, we find no significant correlation between age of exposure and either the Japanese subjects' performance, $r = -.314$, $p > .05$, or the Chinese subjects' performance, $r = -.044$, $p > .05$. Thus, performance on this task does not appear to be influenced by age of exposure. The effect of number of years English had been studied and number of years spent in North America on performance was also examined. There was a (non–significant) moderate negative correlation between the number of years English had been studied by the subjects and performance on each of the experimental conditions across the two groups, $r = -.438$, $p > .05$. These factors were also not significantly correlated within the Japanese group, $r = .117$, $p > .05$, or the Chinese group, $r = .294$, $p > .05$. No significant correlation was discovered between the numbers of years the subjects had spent in North America and performance on this task for the Japanese group, $r = .228$, $p > .05$, the Chinese group, $r = -.043$, $p > .05$, or across these groups, $r = .268$, $p > .05$.

Picture Identification Task

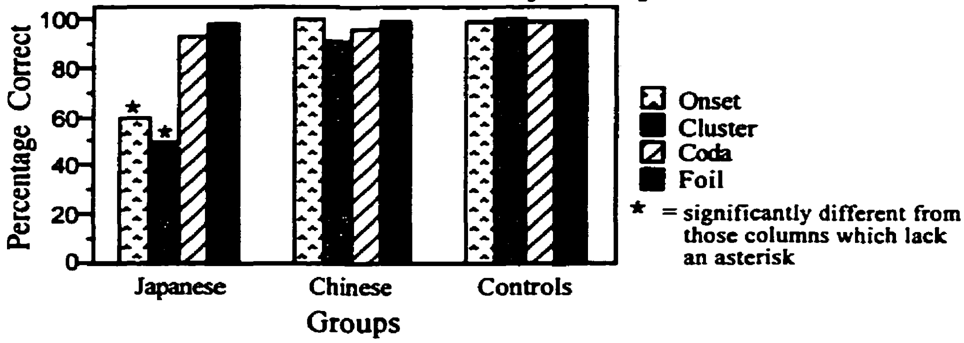
For the picture identification task, selection of the target picture was counted as a correct response and selection of the contrast picture was counted as an error. "Not sure" responses were also considered errors.[23] They were scored this way because the task

---

[23] The reader will notice that "not sure" responses were not included in the tabulation of performance scores on the auditory task. Due to the different natures of the two tasks, it is appropriate to treat "not sure" responses differently in each of the two tasks. In the auditory task, subjects may be

requires the subject to choose one of two pictures based on his or her internal representation of those objects. If the subject's grammar does not distinguish /l/ and /r/, then he or she will be unable to choose the correct picture and "not sure" is an appropriate, albeit incorrect, response. There were, in fact, very few "not sure" responses across subjects, suggesting that if they were not sure they just guessed.

Figure 4
Overall Picture Performance By Group



We will begin with a comparison of the groups' performances, given in Figure 4. There was again no difference between performance of the two test versions and they were, therefore, pooled. A quick comparison of this graph with Figure 3 shows that the overall pattern of performance across groups on this task is very similar to the pattern on the auditory task. Near perfect to perfect performance was attained for all groups on the foils in this task. Thus, as with the auditory task, performance on the /l/–/r/ contrast can be interpreted independently of the task per se. A factorial ANOVA reveals that groups performed significantly differently from one another on the onset condition, $F(2, 27) = 43.74$, $p = .0001$, and on the cluster condition, $F(2, 27) = 41.524$, p. = 0001. Scheffe tests show that the Japanese group performed significantly worse than the Chinese and Control groups. There is no significant difference between the Chinese and Control groups. The performance of the Chinese and Japanese groups are considered separately below.

---

unsure about a particular trial if they misheard one of the stimuli; their L1 or interlanguage grammar will not cause uncertainty. In the picture task, on the other hand, subjects are less likely to be unsure due to some non-linguistic reasons; rather, they will be unsure because their L1 or interlanguage grammar provides them no means to distinguish the two pictures.

Japanese

The mean performance scores of the Japanese group shown in Table 4 were subjected to a repeated measures ANOVA.

Table 4
Japanese Performance Picture Task

Subject	Onset	Cluster	Coda
1	75	41.7	**100**
2	75	58.3	75
3	58.3	50	**100**
4	**91.7**	75	**100**
5	75	**83.3**	**100**
6	16.7	8.3	66.7
7	66.7	75	**100**
8	58.3	33	**91.7**
9	25	50	**91.7**
10	50	33	**100**
Means	59.2%	50.8%	**92.5%**

This ANOVA indicates a significant difference between performance on the three experimental conditions, $F(9, 30) = 41.513$, $p = .0001$. According to Scheffe post-hoc tests ($p<.05$), performance on the coda condition was significantly better than that of either the onset or cluster condition. Performance on the onset condition was not significantly different than performance on the cluster condition (59.2% vs. 50%), both are within the range of chance. Thus, these subjects were unable to accurately distinguish two pictures whose names differed minimally when the contrast was in onset or cluster position. When the contrast was in coda position, however, subjects accurately distinguished the pictures.

This observation is confirmed by an examination of individual scores. The majority of subjects are unable to accurately perform this task for the onset and cluster conditions, yet they are able to perform it when the contrast is in coda position. There are two exceptions to this general finding, though. Subject 4 is able to accurately distinguish /l/ and /r/ in onset position and subject 5 is able to distinguish the contrast when it is part of a consonantal cluster.[24] These exceptions will be discussed below.

Chinese

As a group, Chinese subjects accurately performed this task over ninety percent of the time for all experimental conditions. Mean performance on each of the conditions was not significantly different, as revealed by a repeated measures ANOVA, $F(9, 30) = 2.688$, $p = .0663$. Although performance is slightly lower on the cluster condition, we must consult individual scores given in Table 5 in order to determine whether subjects' performance on these items is above chance.

Table 5
Chinese Performance Picture Task

Subject	Onset	Cluster	Coda
11	**100**	75	66.7
12	**100**	**100**	**100**
13	**100**	**100**	**100**
14	**100**	**91.7**	**100**
15	**100**	66.7	**100**
16	**100**	**91.7**	**91.7**
17	**100**	**91.7**	**100**
18	**100**	**91.7**	**100**
19	**100**	**100**	**100**
20	**100**	**100**	**100**
Means	**100%**	**90.8%**	**95.8%**

The individual scores show that eight of the ten subjects reached criterion on all of the experimental conditions. However, subject 15 did not reach criterion on the cluster condition while subject 12 failed to reach criterion on both the cluster and coda conditions. Although there are two exceptions to the general success of the Chinese with this task, it is important to note the clear distinction between the performance of the

24 Both of these subjects have relatively good control over the pronunciation of the two liquids, as determined by casual observation. It is possible that knowledge of correct articulation has been encoded into their lexical representations and they distinguish lexical representations on this basis.

Japanese subjects who are unable to distinguish /l/ and /r/ in this task and the performance of the Chinese subjects who, on the whole, are able to distinguish this contrast in every word position tested.

As with the Auditory task, the difference between language groups on the Picture task suggests that properties of the subjects' L1 is the key factor in determining performance on the task. Correlations between external factors and performance were computed for age of exposure, number of years English had been studied and number of years spent in North America. We found no significant correlation between age of exposure and performance across groups, $r = .234$, $p > .05$. There was a moderate negative correlation between age of exposure and performance for the Japanese group, $r = -.487$, $p > .05$; however this correlation was also not significant. The Chinese group showed a slight non-significant correlation between these two measures, $r = .163$, $p > .05$. We found a significant moderate correlation between performance on this task and the number of years that English had been studied across the two groups, $r = -.460$, $p < .05$. However, this correlation was *negative*, suggesting that fewer number of years studying English predicts better success. This is no doubt due to the fact that the Chinese subjects (who as a group performed better than the Japanese) had, on average, studied English fewer years. When we factor out the L1s, and consider each group individually, we find no significant correlation between years of English studied and performance for either the Japanese group, $r = -.177$, $p > .05$, or the Chinese group, $r = .058$, $p > .05$. There was also no correlation between the numbers of years the subjects had spent in North America and performance for the Japanese group, $r = .322$, $p > .05$, the Chinese group, $r = .431$, $p > .05$, or across these groups, $r = .385$, $p > .05$.

Comparison of Auditory and Picture Tasks

One of the hypotheses guiding this study is that detection of a contrast is a pre-requisite for the construction of the phonological structure to represent that contrast. In other words, if a subject is unable to accurately perform the auditory task, he or she will be unable to accurately perform the picture task; the converse is also true, if a subject accurately performs the picture task, he or she should be able to accurately perform the auditory task. In order to see whether this hypothesis is borne out by the data, individual subjects' performances on the two tasks was compared. In this case, a statistical test is not appropriate, as it will only tell us whether there is a significant difference between performance *levels* on the two tasks; these tests do not tell us whether subjects were able to accurately perform (i.e., reach the criterion level) one task and not the other.

A comparison of individual scores in Tables 2, 3, 4 and 5 indicates that, for each experimental condition, the majority of subjects (eleven of sixteen) are either able to accurately perform both tasks or unable to accurately perform both tasks. Of the remaining subjects, two reached criterion on the auditory task but failed to reach criterion on the picture task (subjects 7, 11) . The final (somewhat problematic) three subjects reached criterion for one condition of the picture task yet failed to do so for the same condition on the auditory task (subjects 4, 5, 16).

A somewhat surprising finding is that the Japanese speakers' performance on the picture task (roughly 60% correct) is, on the whole, better than their performance on the auditory task (roughly 30% correct). Researchers have found that subjects are generally better at identifying non-native sounds in isolation than they are at discriminating those same sounds side by side (MacKain, Best & Strange, 1981; Borden, Gerber & Milsark, 1983). The auditory task in this study requires discrimination of adjacent phonemes whereas the picture task initially requires identification of a single acoustic signal (and subsequent matching of that signal with the appropriate lexical representation). Thus, the differences measured in this study between the two tasks appear to be consistent with previous findings that one type of task is more difficult than the other.

However, I believe that the observed difference in subjects' performance on the two tasks actually stems from the Japanese speakers' grammar, and the effect of the grammar on the particular task. In particular, according to the hypothesis that the native grammar constrains the perception and acquisition of non-native contrasts, then the expected (i.e., baseline) performance for the two tasks is different. According to the hypothesis that the Japanese grammar funnels the acoustic signal for both /l/ and /r/ into a single native phonemic category, Japanese speakers will always perceive minimal pairs as identical. Thus, we should theoretically expect 0% correct performance for discriminating /l/ and /r/ acoustically. In practice, they are able to correctly discriminate the minimal pairs approximately 30% – perhaps due to variation in duration or amplitude, which were not controlled for in this study. In contrast to the auditory task, the expected performance (given the Japanese grammar) for the picture task is 50% accuracy. Since the Japanese speaker is unable to distinguish /l/ from /r/ in the representation of lexical items (i.e., the same phonological representation is used for both), then when the subject is presented with two pictures and a single verbal cue, he or she simply cannot decide which picture corresponds to the cue and has a 50% chance of choosing the correct picture just by chance. The observed performance, then, 60%, is not significantly different from chance. Thus, if the performance of the Japanese speakers is constrained by their native grammar, as I am suggesting, then the expected performance is different for the auditory task versus the picture task, which is precisely what we found. In fact, the differences in performance between these two types of tasks (i.e., discrimination versus identification) that have been reported in the literature may in fact have nothing to do with the inherent difficulty of a particular task, but rather stem from the effects of the native grammar on the performance of the particular task.

3.3.4. Discussion

In summary, the results from the two tasks demonstrate that Japanese speakers are unable to accurately discriminate /l/ and /r/ *both* acoustically and phonologically, whereas Chinese speakers discriminate this contrast with native-like accuracy in both tasks. Our initial hypothesis was that both Japanese and Chinese learners of English would be unable to acoustically discriminate the non-native /l–r/ contrast, since one of the members of this contrast is a phoneme in the learners' respective L1s. This hypothesis was based on the proposal made in this paper that the elaboration of segmental structure in first language acquisition causes a degradation of the innate capacity to perceive non-native contrasts. The results from the auditory task show that the Japanese speakers are, as predicted, unable to accurately discriminate this non-native contrast; this finding is consistent with the large body of existing research on Japanese perception of /l/ and /r/. In stark contrast to the Japanese speakers, however, the Chinese speakers are able to discriminate these two sounds with native-like accuracy. Thus, our initial hypothesis is only partially borne out by the data.

Before considering possible reasons for the difference between the Japanese and Chinese speakers, let's first examine whether a speaker's L1 syllable structure affects his or her perception and acquisition of non-native phonemes. Recall that both Japanese and Chinese lack consonantal clusters and restrict the segments that can occur in coda position. An apparent contradiction to the observation that Japanese speakers are unable to discriminate /l/ and /r/ is their native-like discrimination of this contrast when it is in coda position; their ability to accurately perform both tasks on the coda condition suggests that word-position may be a factor in the acquisition of a contrast. However, as pointed out above, the coda position of a syllable is different from the other syllabic positions investigated in that it may affect the preceding nucleus. The liquids /l/ and /r/ have a considerable effect on the quality of the preceding vowel, so that words that differ by coda /l/ and /r/ are not true minimal pairs (e.g., *ball* vs. *bar*). It is very likely that the Japanese (and perhaps even the Chinese) subjects are acoustically discriminating these types of pairs on the basis of the vowel quality, rather than on the coda consonant itself.[25] They are, in turn, able to encode this difference in vowel

[25] Since the variation between the /r/– and /l/–colored vowels is relatively subtle (compared to the difference between /l/ and /r/), one might wonder how the L2 learners (especially the Japanese speakers) are able to discriminate them. Given that Japanese and Chinese do not contrast these vowels, and assuming the hypothesis that a speaker's L1 will constrain which sounds he or she accurately perceives, we would expect the L1 grammar to "funnel" the stimuli for the /r/– and /l/–colored vowels into a single L1 perceptual category, as happens for /l/ and /r/, preventing the learners from detecting this subtle difference. However, there is an important difference between the ways in which vowels and consonants are handled by the speech perception mechanism. Speech perception research shows that while consonants are perceived categorically (with little to no sensitivity to within category variation), vowels are not perceived categorically; that is, speakers are as good discriminating variation within vowel categories as they are across different vowel categories (much as musical tones are perceived) (Fry, Abramson, Eimas & Liberman, 1962; Pisoni, 1973; Raphael, 1972). Thus, the perception of

quality into their lexical representations, which enables them to distinguish the pairs phonologically. Thus, the performance of the Japanese speakers on the coda condition is not incompatible with the claim that they cannot discriminate the /l/–/r/ contrast; however, neither the Chinese nor the Japanese results can be used to determine whether speakers of these languages can integrate /l/ and /r/ into the coda position of English words.

The Chinese performance on the cluster condition does, however, demonstrate that L1 syllable structure does not have the same "blocking" effect on acoustic perception that an L1 feature geometry does. Although Chinese lacks consonantal clusters, these speakers can perceive the /l/–/r/ contrast in that position; moreover, they are able to incorporate consonantal clusters in their lexical representation of English words. In contrast, Japanese speakers are unable to discriminate /l/ and /r/ in the cluster position. However, the fact that these speakers are also unable to discriminate the contrast in onset position (a syllabic position permitted in Japanese) indicates that low performance on the cluster condition is an artifact of their inability to discriminate the contrast in general, rather than an effect of their L1 syllable structure. These results suggest that there is a fundamental difference between the acquisition of segmental structure and syllable structure, supporting the theoretical distinction made between projected structure (syllabic) and inherent structure (segmental).[26]

Groups were fairly evenly matched for age of exposure, education and time spent in North America and there were no significant correlations between these external factors and performance. An interesting finding in this regard is that there was no effect for age of exposure. The Critical Period hypothesis, as advanced by Lenneberg (1967), maintains that language acquisition must occur before the onset of puberty in order to be successful. Yet Chinese subjects that were exposed to English

---

vowels is not constrained by one's native phonology (either in the monolingual speaker or the L2 learner), which enables the Japanese and Chinese speakers to discriminate subtle distinctions between non-native vowels. In order to determine definitively whether these speakers are discriminating the coda pairs on the basis of the (/l/– or /r/–colored) vowel quality alone, an additional experiment, in which subjects are asked to discriminate minimal /l–r/ coda pairs whose codas have been chopped off (by computer manipulation of the signal), is required.

[26] Broselow & Park (1995) report that Korean learners can acquire the distinction between long and short vowels (a prosodic property) even though length is not contrastive in their native language.

long after any supposed critical period (e.g., 38 years old) were still able to accurately discriminate the contrast. In contrast, Japanese subjects who had been exposed relatively young (e.g., 10 years old) were not able to discriminate between the two sounds.[27] The differential performance of these two groups must, therefore, stem from their respective L1s.

A possible explanation for the differences between the Japanese and Chinese speakers lies in the more subtle phonological properties of their languages. I argued in §4.2 that a speaker's feature geometry channels the acoustic signal into native phonemic categories, in effect blocking the perception of certain non-native contrasts. The converse of this is that a speaker may be able to perceive a non-native contrast if the *feature* that distinguishes the two segments is present in his or her L1 feature geometry for independent reasons (i.e., despite the fact that feature is not utilized in the representation of the native member of the contrast in question). The feature that distinguishes the lateral approximant /l/ from the central approximant /r/ is [Coronal] (see §3.1). An examination of the Japanese and Chinese phoneme inventories reveals that Chinese (but not Japanese) does, in fact, require [Coronal] in the representation of certain phonemes. These inventories are given in (5) (from Maddieson, 1984 and Vance, 1987).

[27] The results of two Japanese subjects were excluded from analysis. These Japanese speakers were born in the United States to Japanese parents and lived in an English-speaking environment for their first year. Upon return to Japan, both subjects were exposed to English via Sesame Street on television and attended kindergarten at age four where they were also exposed to English. The results of these two subjects are quite remarkable and deserve comment: on both the auditory task and the picture task, they perform like English-natives (despite their non-native grasp of English syntax, as measured by a test of reflexive binding in an independent study). These results provide further support for the claim that if a second language learner is able to acoustically perceive a non-native contrast then he or she will be able to acquire that contrast: even Japanese speakers are able to acquire the /l–r/ contrast if they can hear the contrast (just as Chinese speakers). It is clear that early exposure to English facilitated these two subjects' perception of the English contrast. Moreover, given the fact that the other Japanese speakers were not able to perceive the contrast, despite allophonic exposure in their native language, it is likely that mere exposure to /l/ and /r/ was not sufficient to maintain accurate perception of these sounds. Rather, these facts taken together suggest that the two early-exposure subjects had acquired the feature *coronal* and that this feature is responsible for the maintenance of their perception. Tees & Werker (1984) also report maintenance of perception of non-native Hindi contrasts for adults exposed to Hindi during the first two years of life.

(5) Japanese Inventory[28]

p	t		k	
b	d		g	
		tʃ		
		dʒ		
	s			h
	z			
m	n			
	ɾ			
w		j		

Mandarin Chinese Inventory

p	t		k	
pʰ	tʰ		kʰ	
	ts	ṭs		
	s	ṣ		h
		ẓ		
m	n			
	l			
w	j			

Mandarin contrasts alveolar and retroflex sibilants (e.g., /s/ and /ṣ/). Both of these segments must, therefore, contain [Coronal] in their representations (in order to contrast these two fricatives, either the alveolar /s/ or the retroflex /ṣ/ will be further specified with a Coronal dependent). Crucially, a Chinese speaker will have posited the feature [Coronal] in the elaboration of his or her geometry (independently of the representation of any liquid). Japanese, on the other hand, does not contrast any phonemes within the coronal place of articulation. Hence, the Coronal node will not be present in a Japanese speaker's feature geometry.[29]

According to this explanation, Chinese speakers are able to perceive the /l/–/r/ contrast by virtue of the presence of the feature [Coronal] in their L1 grammar, despite the fact that this feature is not present in the representation of their native liquid /l/.[30] If

[28] While other segments are realized phonetically in Japanese, such as ɸ, they are derived (i.e., occurring in predictable phonological contexts) and do not, therefore, constitute independent phonemes.

[29] Although there are coronal segments in Japanese (e.g. /t/, /s/, /n/), under a theory of Minimally Contrastive Specification (Avery & Rice, 1989), a feature will only be present in a grammar if that feature is required to contrast segments; accordingly, coronal segments in Japanese will be represented with a bare Place node. However, based on palatal prosody in Japanese mimetics, Mester & Itô (1989) argue that Japanese coronal segments are, in fact, represented with the feature [coronal]. This specification is proposed in order to account for the fact that all coronal segments *except* /ɾ/ are palatalized (as are non-coronals). By specifying all coronal sounds, except /ɾ/, with the feature [coronal], the authors explain why /ɾ/ is not a target of this operation. The same facts, however, can be obtained by assuming (as the authors themselves do to explain why /ɾ/ cannot be geminated) that /ɾ/ is not specified for any Place Node at all, whereas coronal segments are specified for a bare Place Node (with no [coronal] feature). Lacking a Place Node, /ɾ/ will never be the target of place assimilation. This specification would also explain why coronals are the preferred target of this operation, with non-coronal becoming palatalized only in the absence of another coronal: lacking Place features, the palatal morpheme has a free place to dock on coronals, whereas the addition of palatalization to the non-coronal segments creates a less-favored complex structure. Thus, the Japanese palatal prosody data do not require that coronals be specified for [coronal] and we can maintain the position that this feature is not manipulated by the Japanese phonology. (Note that whether the representation for /ɾ/ contains a bare Place Node or no Place Node at all does not affect the thesis of this paper)

[30] These results are of course expected under the analysis that Mandarin Chinese contrasts /l/ and /r/. This situation would be analogous to the one described in §4.3, in which each member of the non-native contrast corresponds to segments that are distinct in the speaker's L1. If this is indeed the case,

this is correct, these results support the proposal that it is the feature geometry itself (as opposed to individual segmental representations of phonemes in the inventory) that underlies the boundaries of categorical speech perception.[31] The performance of the Japanese speakers suggests that if a speaker's L1 grammar lacks the feature that distinguishes two segments, he or she will be unable to perceive that contrast.

A second experiment was conducted in order to establish more conclusively whether it is indeed the features (not the phonemes) of the L1 phonology that constrain perception and, thereby, determine which non-native contrasts will be successfully acquired. If it is the presence or absence of the relevant feature in the L1 of the learner which is responsible for cross-language differences in L2 acquisition (as argued above for Japanese and Chinese learners of English), then the status of the contrasting feature in a speaker's L1 should also account for any variation in the acquisition of various non-native contrasts by speakers of the same L1. If it is indeed the featural level that impinges upon L2 acquisition, then we should find instances in which this level of knowledge facilitates acquisition, in addition to instances where it hinders acquisition.

3.4 Experiment 2

3.4.1. Additional properties of Japanese & the contrasts investigated

The results of experiment 1 suggest that the distinctive features manipulated in the L1 grammar are a more important factor in the acquisition of non-native phonemic contrasts than whether the L1 inventory contains either member of the non-native contrast. We

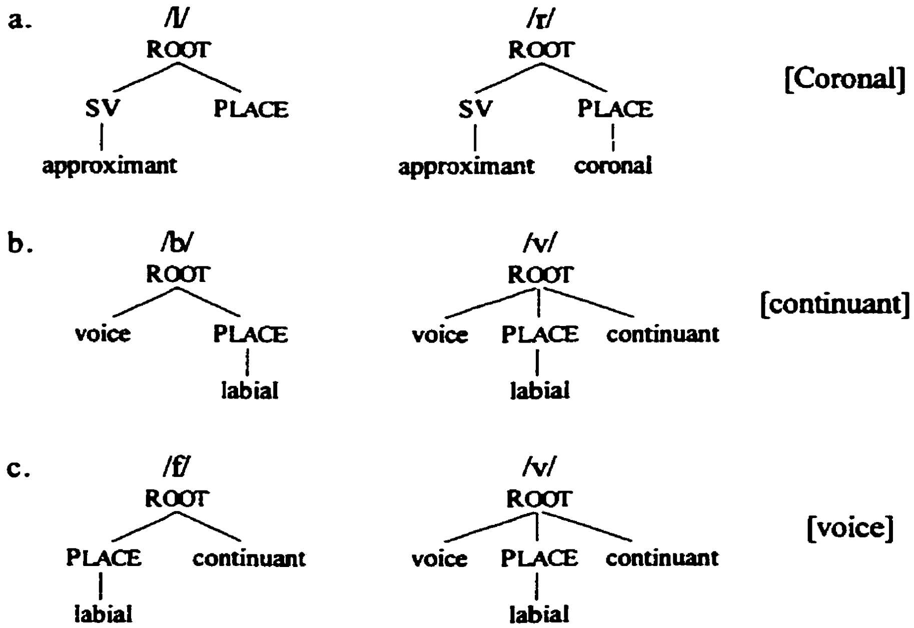
---

then Chinese speakers are able to perceive the English contrast by virtue of its similarity to the native segment /ʐ/; they then substitute /ʐ/ for English /r/ in their lexical representations. There is evidence, however, that is NOT what they are doing. The results from a follow-up study indicate that Chinese speakers can, in fact, distinguish their native "r" (/ʐ/) from the English /r/, both acoustically and phonologically. This would be unexpected if Chinese speakers heard the English /r/ as their native category and merely substituted their native category in lexical representations of English words.

[31] The differential performance of the two language groups also speaks to phonological theory, providing experimental evidence for the representations of /l/ and /r/ assumed in this paper. According to Brown (1995), /l/ and /r/ are distinguished by the presence of the coronal node in the representation of /r/. We can interpret the differential performance of the two groups in terms of the presence of this feature in Chinese and the lack of it in Japanese. However, according to Rice & Avery (1991b), /l/ and /r/ are differentiated not in terms of place features, but by manner features: /r/ contains a vocalic node whereas /l/ does not. Chinese and Japanese do not differ with respect to this feature, thus Rice & Avery's representations incorrectly predict that Chinese and Japanese speakers should perform similarly.

can test this hypothesis by comparing acquisition of non-native contrasts which differ with respect to whether the relevant feature exists in the L1 grammar of a homogeneous group of speakers. This second experiment compares the acquisition of the English /l–r/, /b–v/ and /f–v/ contrasts by Japanese speakers. These contrasts were chosen to test the proposed model of phonological interference because these pairs are not contrastive in Japanese and each causes a differing degree of difficulty for these learners. The internal structure of each contrast under investigation is given in (6); the phonological feature that distinguishes each contrast is given to the right (the superordinate SUPRALARYNGEAL and LARGYNGEAL components are not relevant for this discussion and have been omitted for ease of exposition; the representations for /r/ and /l/ are repeated here from (2)).

(6) *Representations* *Contrasting feature*



These are the representations that the learner must acquire in order to distinguish these phonemes in his or her interlanguage grammar. The important thing to note is that each pair of phonemes is minimally differentiated by the presence of a single phonological feature. Based on the results of experiment 1, our hypothesis in this experiment is that the L2 learner will only acquire those contrasts that are distinguished by a feature that exists (for independent reasons) in his or her native grammar.

Let's consider again the relevant properties of Japanese phonology. From the inventory given above in (5), we can see that each of the contrasts under investigation has a slightly different status, summarized in Table 6.

Table 6
Status of English Contrasts in Japanese

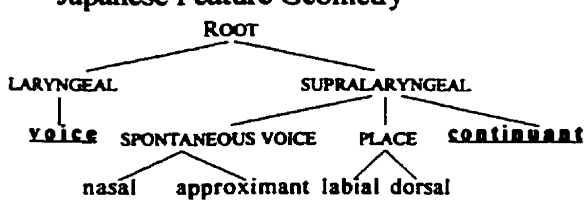
	English Contrasts			
Status in Japanese Inventory	/b/ vs. /v/	/f/ vs. /v/	/r/ vs. /l/	/p/ vs. /b/
is a native phoneme	√			√ √
corresponds to a native phoneme		√	√	
does not correspond to a native phoneme	√	√	√	

Notice that, with respect to the inventory, the status of the /f–v/ and /l–r/ contrasts is similar. If it is the phonemes of the L1 which constrain perception, then we would expect the /l–r/ and /f–v/ contrasts to pattern together in acquisition (in contrast to /b–v/).

The inventory also allows us to determine which phonological features are used contrastively in Japanese; the adult feature geometry in (7) illustrates which features are manipulated.

(7)

Japanese Feature Geometry



Importantly, two of the three non-native contrasts under investigation are distinguished by a feature that is utilized in the Japanese phonological system. The feature [continuant] distinguishes /b/ from /v/ and is independently required in the Japanese grammar to differentiate native /stop–continuant/ contrasts (e.g., /t–s/). The other non-native contrast, /f–v/, is distinguished by the feature [voice], which is required in the Japanese grammar to represent native voicing contrasts (e.g., /p–b/). As pointed out in the first experiment, the feature which distinguishes /l/ from /r/, [Coronal], is not manipulated in the Japanese grammar since there are no contrasts within the coronal place of articulation.

To summarize, then, this second experiment was designed to discover whether there is a difference between the acquisition of a non-native contrast which is distinguished by a feature NOT contained in the L2 learner's native grammar (i.e., /l–r/) and the acquisition of a non-native contrast which is distinguished by a feature that IS contained in the L2 learner's native (i.e., /b–v/ and /f–v/). In the previous experiment, the effect of the status of the relevant feature on acquisition was examined across groups; the present experiment examines this variable within a single language group, enabling a more direct assessment of the impact of L1 features on non-native phoneme acquisition. According to the model of phonological interference developed in this paper, our hypothesis is that the Japanese speakers will accurately perceive the /b-v/ and /f-v/ contrasts, but not the /l-r/ contrast. Moreover, they should perceive these two non-native contrasts and the native /p-b/ contrast equally well.[32] As a result, successful acquisition of these two non-native contrasts, again to the exclusion of /l-r/, should follow.

3.4.2. Methodology

Subjects

Thirty adults with no evidence of hearing loss participated in experiment 2. The experimental group consisted of fifteen Japanese speakers who had learned English as their only second language. Each of these subjects was raised in Japan and had come to

[32] Uniform performance is not predicted if the aspect of the native grammar responsible for filtering non-native sounds is the phonemic representations themselves (rather than the features). Recall from table 1, that the members of the /b-v/ and /f-v/ contrasts have a different status *vis à vis* the Japanese phoneme inventory. Hence, if the phonemic representations constrain perception, we might expect differential perception of these two pairs, since the segment /b/, but not /f/, occurs in Japanese.

North America to study in an undergraduate or graduate program at McGill University in Montreal, Canada. These subjects were between twenty and thirty-two years old at the time of testing, which was conducted in Canada. The reported age of first exposure to English for these subjects ranged from four years old to thirteen years old. Each subject had studied English in school in Japan for a minimum of six years, up to a maximum of ten years. The length of time these subjects had been continuously living in North America ranged from one month to five years. The control group consisted of fifteen native monolingual English speakers, who ranged in age from 15 years to 54 years. This background information is summarized in Table 7. All subjects were paid for their participation.

Table 7
Subject Information

Group	Mean Age at Testing	Mean Age of Exposure	Mean Years Studied	Mean Years in N. America
Japanese	24.5	9	8	3.5
Controls	25	—	—	—

Tasks & Materials

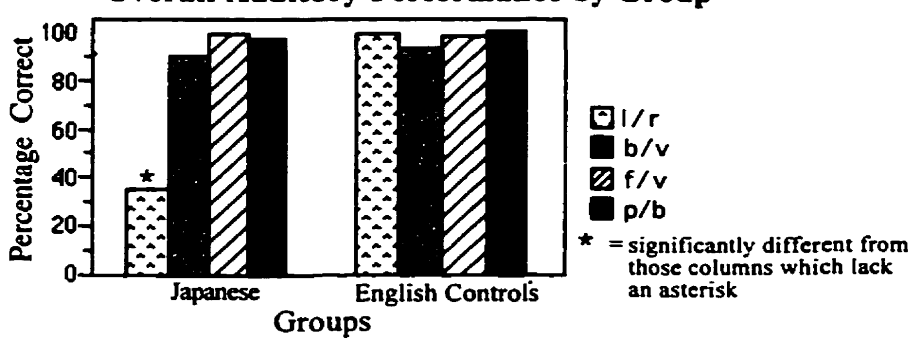
The tasks used in experiment 2 were the same as those used in experiment 1: an AX Discrimination task was used to measure auditory perception and a Forced Choice Picture Selection task was used to assess phonological competence. As different contrasts were investigated in this experiment, the materials used differed slightly from those used in experiment 1. In particular, in the present experiment, the non-native contrasts in question were only examined in the onset position of the syllable (e.g., *bat / vat*). As in experiment 1, two types of foils were included in the test materials: the native /p–b/ contrast was included for comparison with the non-native contrasts, and identical pairs of words were included in the AX task to detect any response bias towards responding that all of the pairs were different. The items used in the test were natural tokens of real English monosyllabic words; see Appendix C for a complete list of stimuli used. These tokens were spoken by a male speaker of American English (the same person as in experiment 1) and recorded so that the stimuli were identical for every subject. Six pairs of each of the four contrasts were included, for a total of sixty items in the AX task and forty-eight items in the Picture task. These materials were counterbalanced for order and position of presentation within a single trial, as was the order of the trials within the entire test. The testing procedure was identical to that in experiment 1.

3.4.3. Results

Auditory Discrimination Task

For the auditory discrimination task, a response that the two words in the minimal pair were "different" was counted as correct and a "same" response was counted as an error. As before, performance on the identical foil pairs was not included since accurate performance on these items would be consistent with the subject's L1 and would not reveal properties of his or her interlanguage grammar. Performance scores on each of the contrasts were tabulated separately for statistical analysis. Figure 5 reports the mean performance scores of both groups on each of the contrasts.

Figure 5
Overall Auditory Performance by Group



From the Japanese subjects' near perfect performance on the control items (i.e., native /p–b/ pairs), it is clear that the task itself does not pose any difficulty for the learners. Thus, performance on the non-native contrasts can be interpreted to reflect properties of the speakers' interlanguage grammar. The Japanese speakers were significantly poorer than the English controls at discriminating the /l-r/ contrast [t (28) = –16.16, p =.0001]. Yet, there was no statistical difference between the two groups in their ability to perceive the other contrasts; the Japanese speakers discriminated each of these English contrasts as accurately as the native controls [/b–v/ contrast: t (28) = –1.28, p = .21; /f–v/ contrast: t (28) = 1.87, p = .08; /p–b/ contrast: t (28) = –1.46, p = .15].

In order to evaluate performance on each contrast relative to the other contrasts, additional analyses were carried out separately on the two groups. Performance scores for the Japanese group are given in Table 8.

Table 8
Japanese Performance Auditory Task

Subject	l/r	b/v	f/v	p/b
1	41.7	**100**	**100**	**100**
2	58.3	**91.7**	**100**	**91.7**
3	8.3	**100**	**100**	**91.7**
4	25	**91.7**	**100**	**100**
5	33.3	**100**	**100**	**100**
6	50	**100**	**91.7**	**91.7**
7	41.7	**83.3**	**100**	**91.7**
8	66.7	**100**	**83.3**	**100**
9	33.3	75	**100**	**100**
10	0	75	**100**	**91.7**
11	8.3	**91.7**	**91.7**	**100**
12	66.7	**100**	**100**	**100**
13	25	**83.3**	**100**	**91.7**
14	33.3	75	**83.3**	**91.7**
15	41.7	**83.3**	**91.7**	**83.3**
Means	35.6%	**90%**	**96.1%**	**95%**

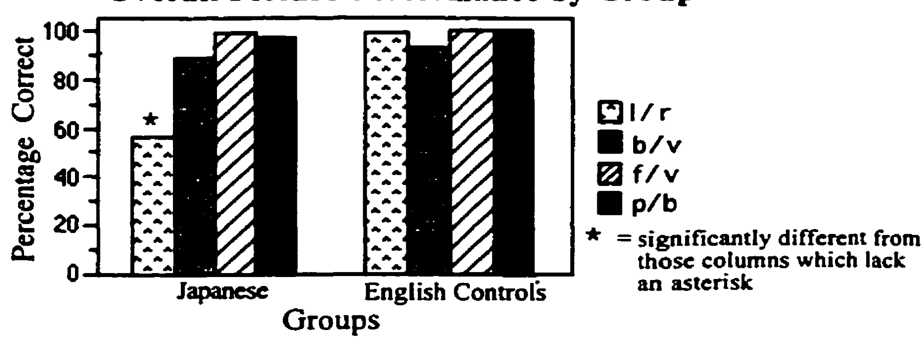
A repeated measures ANOVA on the mean scores shows a significant difference between the experimental contrasts, F (14, 45) = 119.85, p = .0001. Post-hoc Scheffe tests (p<.05) indicate that performance on the /l-r/ contrast is significantly worse than their performance on the other three contrasts; however, performance on the other three contrasts (/b-v/, /f-v/, /p-b/) is not significantly different from one another. This pattern of performance is confirmed by an examination of the individual scores. As in experiment 1, a score of 83.3% (ten out of twelve trials) has been set as criterion for successful performance on a given contrast (p<.05); bolded typeface indicates that this criterion has been reached. All of the subjects reached criterion on the /f–v/ contrast and all but three did on the /b–v/ contrast; however, no subject reached criterion on the /l–r/ contrast.

We must also examine the performance of the control group, as any weaknesses in the experimental materials will by revealed in the performance of the native speakers. Moreover, we can only interpret the performance of the experimental group relative to the control group. Although the native controls' performance on the /b-v/ contrast appears, in the graph, to be depressed relative to the other contrasts, a repeated measures ANOVA reveals no significant difference between the four contrasts, F (14, 45) = 1.44, p = .25. Thus, we can regard the Japanese speakers' performance, in light of the native speaker data, as an accurate reflection of their perceptual capabilities.

Picture Identification Task

For the picture identification task, selection of the target picture was counted as a correct response and selection of the contrast picture was counted as an error. The groups' overall performances are compared in Figure 6. Near perfect performance was attained by the Japanese group on the control items in this task. Thus, performance on the non-native contrasts can be interpreted independently of the task per se.

Figure 6
Overall Picture Performance by Group



As in the auditory task, the Japanese speakers were significantly poorer than the English controls at differentiating the /l-r/ contrast, t (28) = –9.73, p =.0001. There was, however, no statistical difference between the two groups in their ability to distinguish the other contrasts [/b–v/ contrast, t (28) = –1.8, p = .08; /f–v/ contrast, t (28) = –.32, p = .75; /p–b/ contrast, t (28) = –1.27, p = .22].

Let us now compare performance on each of the contrasts relative to each other. A repeated measures ANOVA on the Japanese scores indicates a significant difference between performance on the four experimental conditions, F (14, 45) = 57.65, p = .0001. According to post-hoc procedures (p<.05), the Japanese speakers' performance on the /l–r/ contrast was significantly worse than performance on the other three contrasts, while performance on the non-native /b-v/ and /f–v/ contrasts was not significantly different from each other, nor was it different from performance on the native /p-b/ contrast. This pattern of performance is confirmed by an examination of the individual scores in Table 9.

Table 9
Japanese Performance Picture Task

Subject	l/r	b/v	f/v	p/b
1	75	100	100	100
2	75	75	100	91.7
3	58.3	83.3	100	91.7
4	58.3	83.3	100	91.7
5	75	100	100	100
6	16.7	100	100	83.3
7	66.7	83.3	100	91.7
8	58.3	91.7	100	83.3
9	25	83.3	83.3	100
10	50	75	91.7	91.7
11	41.7	100	100	100
12	58.3	83.3	100	100
13	33.3	100	91.7	100
14	50	83.3	83.3	100
15	50	91.7	100	83.3
Means	52.7%	88.9%	96.7%	93.9%

Looking at the native speaker controls, their mean performance on each of the contrasts was not significantly different from each other, F (14, 45) = 1.56, p = .21. Thus, as with the auditory task, the performance of the Japanese can be taken to accurately reflect their underlying phonological competence.

3.4.4. Discussion

The hypothesis guiding this experiment was that perception of non-native contrasts is constrained by the phonological features manipulated in the native grammar of the learner, not by the phonemes contained in the L1 inventory. This led us to predict that Japanese learners of English would accurately perceive the /b-v/ and /f-v/ contrasts, since these two pairs are differentiated by features already present in the Japanese grammar, but that accurate perception of the /l-r/ contrast would be blocked by the absence of the relevant feature from the Japanese grammar. Each of these predictions was borne out by the data. Thus, our findings regarding the perception of the /l–r/ contrast replicate the results we obtained in experiment 1: the Japanese speakers are unable to discriminate /l/ from /r/, perceiving them, instead, as a single category. Once again, their inability to perceive /l/ and /r/ as distinct phonemes can also be understood as a direct consequence of the influence of the native grammar on the operation of the speech perception mechanism. But it is the Japanese speakers' performance on the /b-v/ and /f-v/ contrast which is most striking: not only do the Japanese speakers perceive the non-native /b-v/ and /f-v/ contrasts with native-like accuracy, as predicted, but they perceive them equally well. This is what we would expect, given that they are both distinguished by a feature in the Japanese grammar.

We therefore have strong support for the proposal that the presence of the features [continuant] and [voice] in the Japanese grammar permits these speakers to perceive the contrast between /b/ and /v/ and between /f/ and /v/, respectively. This means that, despite lack of acoustic, phonetic or phonemic experience with a non-native contrast, a speaker's experience perceiving native sounds along a particular acoustic dimension permits him or her to accurately discriminate any non-native sounds that differ along this dimension (by virtue of the fact that the phonological feature that underlies the acoustic dimension exists in his or her grammar, and serves to sort the incoming acoustic signal).[33] According to this position, there is no difference between

[33] Note that this is true regardless of the actual phonetic realization of a particular contrast. Take, for example, voicing contrasts: although languages may vary as to how they choose to acoustically

perception of native sounds and non-native sounds that vary along the *same* dimension. The Japanese results support this position: there was no difference between their ability to accurately perceive the three non-native contrasts and their ability to accurately perceive the native Japanese contrast, suggesting that perception of non-native sounds operates in the same manner as perception of native sounds.

Turning to the acquisition of new phonological structure, we predicted that Japanese learners would successfully acquire the /b-v/ and /f-v/ contrasts, but would fail to acquire the /l-r/ contrast (since acquisition of feature geometric representations is dependent upon accurate perception of that phonemic contrast). These predictions, too, were confirmed by the data. The Japanese learners successfully acquired only those non-native contrasts which they accurately perceived. When shown two pictures that constituted a minimal /l–r/ pair (e.g., *rake*, *lake*), the Japanese subjects were unable to correctly choose the one that corresponded to the verbal cue; from this we can infer that /l/ and /r/ are not differentiated in their interlanguage grammars. Yet these subjects performed this task with native-like accuracy when the pair of pictures differed by /b-v/ (e.g., *boat*, *vote*) or /f-v/ (e.g., *fan*, *van*).[34] In other words, the phonological structure that represents the /b-v/ and /f-v/ contrasts has successfully been acquired by these learners; using the features contained in the L1 grammar, new representations for the non-native segments have been constructed.

To summarize, then, we have found a difference in the Japanese speakers' ability to perceive non-native contrasts, depending on whether the feature that distinguishes a given contrast exists in their grammar: the /l-r/ contrast (whose contrasting feature is absent from the Japanese grammar) is not accurately perceived, whereas the /b-v/ and /f-v/ contrasts (whose contrasting features are contained in the

---

realize the voicing contrast (i.e. actual VOT's may vary), since the same phonological feature underlies this contrast (i.e. [voice]), the claim is that speakers whose native language exploits this feature will be able to perceive all non-native voicing contrasts.

[34] This results is even more surprising given the tendency of many Japanese learners to substitute /b/ for /v/ in production. But, as pointed out above, there is a well-known dissociation between comprehension and production skills, with comprehension being a more accurate reflection of the speaker's phonological knowledge.

Japanese grammar) are accurately perceived, in a native-like manner.[35] We have also found that when the requisite conditions for successful acquisition are met (i.e., sufficient intake to the acquisition device), new phonological representations can, in fact, be constructed.

3.5 General Discussion & Conclusions

We began this study by posing three related questions: How do we account for the failure of L2 learners to acquire (phonological) properties of the target language provided by UG?, Which aspect(s) of the L1 grammar impinges upon L2 (phonological) acquisition?, and When the necessary conditions for successful acquisition are met, can L2 learners acquire novel (phonological) representations? Let's consider now how the model of phonological interference developed in this paper and our experimental findings contribute to these SLA issues. First, this study demonstrates that we cannot necessarily take learners' failure to acquire UG-determined properties of their L2 as evidence that UG does not operate in L2 acquisition. Before concluding that UG is inaccessible, we must first determine (experimentally, if possible) that the conditions necessary for successful acquisition have indeed been met.

Regarding phonological acquisition, we know from first language acquisition research that the development of segmental structure involves the interaction of UG and the learner's detection of phonemic contrasts in the input. Thus, successful acquisition of novel phonemes by L2 learners depends not only on the availability of UG, but, crucially, on accurate perception of the input. By demonstrating that some L2 learners

---

[35] One might be tempted to surmise that the Japanese speakers are unable to discriminate the /l–r/ contrast precisely because they have allophonic experience with these segments: given the allophonic variation in the native language, we could imagine that these speakers have been "trained" to ignore these variations (recognizing them both as instantiations of the same underlying phoneme). Under this account, the allophonic exposure, rather than the absence of the contrasting feature, would be responsible for their lack of perceptual sensitivity to this contrast. However, this allophonic explanation cannot be correct because /b/ also varies allophonically in Japanese, with the voiced bilabial fricative [β], which shares acoustic and phonological properties with English /v/. According to Kawakami (1977:32), the phoneme /b/ is realized as a plosive word-initially, but is often realized as a voiced bilabial fricative word-internally (compare [bareru] "be revealed" with [aβareru] "rampage"). Thus, if allophonic variation were the cause of the learners' inability to accurately perceive certain non-native contrasts, we would expect perception of the /l–r/ and /b–v/ contrasts to be similarly impaired.

do not perceive the L2 input correctly (in fact, precisely those learners who fail to acquire the new segmental representations), this research indicates that the inability of some L2 learners to acquire novel phonemic contrasts is due to factors other than the inaccessibility of UG. Research on phonological acquisition provides us an opportunity to directly ascertain whether sufficient intake is being received by the learner, and to demonstrate that when there is not sufficient intake, acquisition is not triggered. Thus, this study provides support for the recent trend in SLA theories to attribute L2 learners' failure to acquire certain properties of the L2 to factors other than the inaccessibility of UG.

If the failure to acquire L2 phonological knowledge is a result of insufficient intake to the acquisition device, then what determines this intake? In other words, what prevents the learner from receiving the necessary intake? Here, our model of phonological interference and our experimental findings indicate that the native grammar of a speaker constrains the actual intake to the acquisition device, suggesting that not only may properties of a speaker's L1 be transferred into L2 acquisition (as has been shown for various syntactic constructions, see White, 1989a, for a review), but a speaker's L1 may actually block accurate perception of the L2 input (which may, in fact, turn out to be the cause of transfer in general). This blocking effect of the L1 can have dire consequences for the L2 acquisition of phonology, as demonstrated by the Japanese speakers' inability to acquire the /l–r/ contrast.

These experimental data help establish more precisely *what* the role of the L1 grammar is in L2 acquisition: it does not simply serve as a resource of "building" material for the interlanguage grammar; rather it intervenes between the L2 input and the acquisition device. The native grammar's effect on L2 acquisition is simply an artifact of how it operates in the monolingual speaker – mapping the fuzzy, overlapping speech stream onto discrete abstract linguistic representations. Those variations in the acoustic signal that do not contribute to differences in meaning are simply not perceived by the listener. In the case of L2 phonological acquisition, variation in the acoustic signal which is filtered out by the native phonological system (i.e., is treated as *intra*-category variation) may, in fact, contribute to differences in meaning in the L2 (i.e., actually constitute *inter*-category variation). Thus, the speech perception system will try to map all input onto existing representations; acquisition will only be triggered when (and in those cases where) the L2 input cannot be accommodated by L1 structures (Schwartz, 1996).

This leads us to the second questions posed, namely which aspect of the grammar (or level of representation) impinges upon L2 acquisition. I have proposed that the monotonic acquisition of feature geometric structure by children reduces their perceptual sensitivity to particular non-native contrasts and that this adult feature geometry continues to mediate between the acoustic signal and the linguistic processor in adult speech perception. As a result, once a learner has acquired his or her native feature geometry, all speech sounds (native and non-native) will be perceived in terms of the features exploited by that particular language.[36] Thus, although the L2 input will be perceived *in terms of* the L1 phonemic categories, it is, in fact, the feature geometry that maps the acoustic signal onto those existing categories.

It appears, then, that it is generally not the L1 representations themselves, but the components of those representations which effect (i.e., cause) L2 transfer. SLA researchers have pointed out the fact that some aspects of the L2 seem to be acquired with little or no L1 influence, whereas other aspects of the L2 are heavily affected by properties of the L1 grammar (Schwartz, 1996). This would also seem to be the case with respect to phonological acquisition: some phonemic contrasts are impossible for certain L2 learners to acquire whereas others are acquired easily. However, looking at the featural level of phonological knowledge, we see that there is no "partial influence". The L1 feature geometry operates uniformly in mapping the L2 input to L1 categories; however, in those cases where the L1 does not appear to influence acquisition, it is simply the case that the L1 geometry facilitates perception and acquisition, whereas in

[36] It is possible that the findings in Best et al. (1988) with Zulu clicks can also be explained along these lines. The clicks tested differ with respect to places of articulation (e.g. alveolar, palatoalveolar and lateral) that are employed by English speakers. It is possible that these speakers discriminated these non-native clicks using L1 features, just as the Japanese speakers discriminated the /b–v/ and /f–v/ contrasts (experiment 2), and the Chinese speakers discriminated /l/ and /r/ (experiment 1).

those cases where the native grammar does appear to have an influence, the L1 geometry hinders perception and, thus, prevents acquisition. Yet, the operation of the L1 geometry in perception (i.e., L1 interference) is the same for all non-native sounds.

The theory of phonological interference developed in this paper helps us to better understand *why* and *how* the L1 exerts an influence on the acquisition of a second phonological system. By isolating those aspects of the L1 grammar which are responsible for this influence (i.e., the phonological features) as well as providing an explanation of the mechanisms which exert this influence (i.e., the operation of speech perception in the monolingual), we can not only explain differences in a learner's ability to acquire various non-native contrasts, we can also account for the varying abilities of learners with different L1's to acquire the same non-native contrast.

Finally, we have evidence that when all the necessary conditions are met, L2 learners can successfully acquire UG-determined properties of the target language: the Chinese results on the /l–r/ contrast and the Japanese results on the /b–v/ and /f–v/ contrasts demonstrate that if L2 learners are able to auditorily perceive a contrast not present in their L1, then they are able to acquire that contrast. This suggests that the mechanism for constructing novel segmental representations (which is arguably part of UG) is still operative in L2 acquisition.[37] Of course, the Chinese speakers were intermediate or advanced speakers of English and all of them had acquired the /l–r/ contrast at the time of testing; the Japanese speakers were likewise relatively advanced and provided little evidence for stages of acquisition. We still need to investigate whether there is a developmental progression in the acquisition of segmental structure, by testing beginners in addition to the more advanced speakers, as well as documenting the acquisition process longitudinally. In the beginning, we should find a stage where

---

[37] Of course, in order to establish more conclusively that acquisition of these new representations is constrained by UG, we also need to provide evidence as to the internal structure of these newly acquired phonemes; that is, whether the learners have indeed constructed feature geometric representations and that those representations respect the hierarchical organization of Feature Geometry in UG. Evidence for the internal structure of segments usually comes from the behavior or a given segment in phonological processes (i.e., whether it is a trigger, target or transparent in phonological operations such as assimilation). Thus, in the case of L2 learners, we would need to show that the acquired segment participated in phonological processes in the interlanguage in the same way it does in the grammar of native speakers. Note that this sort of evidence would require that the learner had also acquired the relevant phonological rule.

the learner receives adequate intake (i.e., accurately perceive two sounds), but has not yet acquired the phonological representations for those segments.

Preface to Chapter 4

The investigation of L2 phoneme acquisition in the preceding chapter establishes two important points. First, this study demonstrates that the failure of L2 learners to acquire properties of the L2 cannot necessarily be taken as evidence that UG is inaccessible; our experimental results indicate that L2 learners' failure to acquire novel segmental representations can be attributed to their inability to perceive the contrast in the input, rather than to the inaccessibility of UG. When learners do detect a contrast in the input, new representations are successfully acquired for those segments, suggesting that the mechanism for constructing segmental representations is still operative in L2 acquisition. Thus, this research provides support for the position that UG is accessible in L2 acquisition, but that the L1 grammar impinges somehow on the L2 acquisition process, preventing its full operation in certain circumstances.

Secondly, the preceding study establishes how the L1 grammar affects the L2 acquisition process and which aspects of the L1 grammar are responsible for this effect: the L1 feature-geometry mediates between the acoustic signal and the linguistic system, filtering out the phonetic information in the signal that does not contribute to meaning in the L1 and, thus, restricting the learner's sensitivity to the full range of phonetic information contained in the L2 input. In this way, the L1 grammar actually disrupts the process of acquisition, by limiting which properties of the L2 input the learner will be sensitive to and, therefore, which properties of the L2 phonological system the learner will successfully acquire. Our evidence for this came from Japanese speakers' inability to distinguish /l/ and /r/, either auditorily or phonologically: both segments are perceived as a single category and both are, therefore, represented in the interlanguage grammar by the same feature-geometric structure. This result follows straightforwardly from the theory of L1 phonological interference developed in the previous chapter and the representations proposed in Chapter 1. According to the arguments and data presented there, the feature-geometric representations for /l/ and /r/ are distinguished by the feature [coronal]. This, combined with our theory of L1 phoneme acquisition, accurately predicts that [coronal] would figure importantly in the acquisition of these two segments and that any psycholinguistic effects of these acquired representations would result from their coronality. Lacking the phonological feature [coronal], the feature geometry in the Japanese speakers' mental grammar funnels the acoustic signals for /l/ and /r/ into a single perceptual category, whereas the [coronal] feature present in the Chinese speakers' mental grammar functions to sort the two acoustic signals into distinct perceptual categories. Their detection of this contrast in the L2 input triggers the acquisition of novel segmental representations in which the L1 features are combined in new ways to yield L2 contrasts.

While the experimental data reported and discussed in the previous chapter strongly suggest that it is the features utilized in the *phonological* system (i.e. underlying features) that constrain speech perception and, thereby, impinge upon L2 acquisition, an alternative analysis, which relies on properties of the *phonetic* system, seems plausible and, thus, deserves a closer look. Consider the /l–r/ contrast. I have asserted in Chapter 1 that these segments are distinguished phonologically by the feature [coronal]; though it is argued that there is no *phonological* feature [lateral], a *phonetic* feature {lateral} will be inserted into the representation for /l/ in the phonetic component of the system, ensuring that the appropriate articulation is executed to produce a lateral segment. Chinese and Japanese differ with respect to this phonetic feature {lateral}, much as they differ with respect to the underlying feature [coronal]: both of these languages contain an approximant in their inventory; however, this approximant is realized as a lateral in Chinese and, therefore, this language utilizes the phonetic feature {lateral}. Japanese, on the other hand, does not utilize this phonetic feature since its approximant is not realized as a lateral.[1]

[1] The underlying segment /ɾ/ in Japanese is freely realized as [ɾ], [ɹ], [ɖ], [l] (International Phonetic Association, 1979; Vance, 1987). For our purposes, the important point is that {lateral} does not play a role in the phonetic system of Japanese.

Suppose, for a moment, that it is the phonetic (rather than phonological) features of the system that impinge upon speech perception. Since Chinese and Japanese differ with respect to the phonetic feature {lateral}, it could be argued that the differing abilities of Chinese and Japanese speakers to perceive the /l–r/ contrast result from the fact that Chinese possesses {lateral} and Japanese does not; in this analysis, the existence of the appropriate *phonetic* feature in the L1 grammar would enable an L2 learner to perceive a given non-native contrast. If this analysis were correct, then our L2 perceptual data could not be taken as support for the claim that feature-geometric (i.e., underlying) representations have psycholinguistic consequences for the perception of speech sounds or that the L1 feature geometry has consequences for the organization of the speech perception system. Thus, although this "phonetic-feature" analysis lacks a theoretical underpinning (i.e., as yet we have no theory of how the influence of the L1 phonetic system might arise), it appears to provide a plausible account for the cross-language differences in L2 perception (at least for the /l–r/ contrast), and must, therefore, be empirically tested in order to be ruled out.

Taking the findings and conclusions of the previous chapter as its point of departure, the following chapter extends the line of inquiry developed in the prior study by investigating an expanded set of L1 groups and additional non-native contrasts.[2] Incorporating Korean speakers into our experimental paradigm (in addition to Japanese and Chinese speakers) enables us to demonstrate conclusively that it is the organization of L1 phonological features (not phonetic features) that guides the mapping of the L2 acoustic signal onto existing phonemic categories. Korean is similar to both Chinese and Japanese in that its inventory contains a single approximant segment. However, it is more similar to Chinese with respect to certain phonetic properties and, conversely, more similar to Japanese with respect to certain phonological properties; this three-way comparison will allow us to tease apart the putative effects of phonetic and phonological

---

[2] Note that, as each of these chapters is an independent paper, the theoretical background and theory of L1 phonological interference presented in the preceding paper is outlined again in Chapter 4 in order to provide an appropriate and sufficient context for an extension of the previous research. In addition, a few of the figures from Chapter 3 are reproduced in Chapter 4 to illustrate the model of L1 phonological interference advanced here; experiment 2 from the preceding chapter is also summarized, as experiment 1, to establish a point of departure for the subsequent experimental investigations.

features on the perception of L2 sounds. Like Chinese, the approximant in Korean is realized as a lateral[3]; thus, both languages utilize the phonetic feature {lateral}. In contrast, Korean is more similar to Japanese in its underlying featural properties: both languages lack the phonological feature [coronal]. Thus, the "phonetic-feature" analysis predicts that the Korean speakers, like Chinese speakers, will accurately perceive /l/ and /r/ segments as distinct, due to the presence of {lateral} in their L1 grammars. In contrast, the model of L1 interference developed in the previous chapter predicts that the Korean speakers, like Japanese speakers, should be unable to distinguish /l/ and /r/ perceptually, due to absence of the feature [coronal] from both L1 grammars. The experimental results show that Korean speakers are unable to perceive a contrast between /l/ and /r/, indicating that it is indeed the organization of L1 *phonological* features that governs the perceptual mapping of the acoustic signal into phonemic categories.

This chapter provides further experimental evidence for this position by comparing learners'' acquisition of several different English contrasts. Japanese, Chinese and Korean speakers are correctly predicted to accurately perceive only those English contrasts that are distinguished by a phonological feature present in their L1 grammars. Thus, not only is the theory of L1 phonological interference espoused here shown to account for cross-language differences in the ability to perceive and acquire a given non-native contrast (i.e., Japanese vs. Korean vs. Chinese speakers' ability to acquire, for example, the /l–r contrast), it is also shown to account for the varying degrees of difficulty that speakers of a single L1 have in acquiring a range of non-native contrasts (Japanese or Korean or Chinese speakers' ability to acquire, for example, the /l–r/ vs. the /b–v/ vs. the /f–v/ contrasts). In addition to comparing different L1's and different non-native contrasts, this study also examines whether the influence of the L1 grammar on speech perception and phoneme acquisition changes as the learner progresses and suggests a course of development for L2 phonemic categories. This

---

[3] The surface realization of the underlying lateral approximant varies distributionally between an apical flap [ɾ] intervocalically and a lateral approximant elsewhere (Jung, 1962).

additional exploration enables us to place L2 phoneme acquisition research into the context of recent trends in SLA theory and provides some answers to the questions that are currently central to the field of second language acquisition.

4 The interelation between speech perception and phonological acquisition from infant to adult

4.0 Introduction

The acquisition of a second language (L2) is clearly *somehow* different from that of a first language (L1): adult second language learners rarely (if ever) achieve the same native competence that children do learning their first language and, conversely, children never experience the difficulties that L2 learners do.[1] This disparity between L2 and L1 acquisition is perhaps most apparent with respect to the acquisition of a second phonological system. Whereas children consistently achieve native competence across the full range of subtle and complex phonological properties of their language, second language learners often have extraordinary difficulty mastering the pronunciation and intonation patterns of their L2, typically retaining an accent even when their syntactic knowledge of the L2 is quite native-like. This lack of success is often taken as evidence that UG does not operate in second language acquisition. But, perhaps there is another explanation. As White (1989a) points out, other factors, in addition to UG, are necessary for successful first and, presumably, second language acquisition (for example, sufficient input and various learning mechanisms). An intriguing line of research suggests that the failure of some L2 learners to attain a native-like competence is attributable to these other factors, rather than to the non-operation of UG.

---

[1] Although the term "second language acquisition" (SLA) technically refers to the acquisition of a second language by either an adult or a child, it is typically used to denote acquisition by post-pubescent learners. In this paper, we will only consider SLA by adults; however, the claims made here may be extended to L2 learners of all ages.

One such factor that distinguishes second language acquisition from first language acquisition is the fact that the second language learner comes to the task of acquisition already knowing a language. Most current theories of second language acquisition do, in fact, assume that the native language of the learner plays a role in acquisition. Although researchers generally agree that the learner's existing linguistic knowledge exerts some influence on the acquisition process, there is considerable debate as to precisely *what* role the native language plays (e.g. Bley-Vroman's *Fundamental Difference Hypothesis*, 1989, versus White's *Transfer Hypothesis*, 1988; see also papers in Schwartz & Eubank, 1996, on the L2 initial state). Moreover, existing research suggests that the influence of the native grammar is not absolute: some aspects of the L1 seem to prevent successful acquisition of particular L2 structures, whereas other properties of the L2 are acquired with little or no interference from the native grammar (Schwartz, in press). The challenge for second language theory now is to provide a principled explanation for the presence or absence of L1 influence, that is, what determines "partial influence".

The position that the native language plays a role in the acquisition process is based largely on research that assesses the syntactic knowledge of L2 learners. However, a considerable amount of research on the acquisition of phonological knowledge by second language learners has been conducted outside the generative theoretical framework, in particular on the acquisition of non-native contrasts (Briere, 1966; Flege, 1981, 1991; Wode, 1978, 1992). Like the syntactic research, this earlier phonological research demonstrates that the native language exerts a substantial influence on the acquisition of a second language. It is now widely accepted among L2 phonology researchers that non-native (L2) sounds are perceived in terms of native (L1) phonemic categories. As Flege (1981:448) states, "the tendency by mature speakers to interpret sounds occurring in a foreign language in terms of sounds found in their native language may be a more important cause of foreign accent than any limitation on phonetic learning imposed by neurophysiological maturation." Yet, although this previous phonological research has addressed the question of *whether* the native language plays a role, it has not attempted to answer the question of *why* the native language influences L2 acquisition, nor has it formally articulated the mechanisms by which the native grammar influences this acquisition. More recently, though, researchers working within the generative framework have begun to examine the role of the learner's native language in the acquisition of phonological knowledge (e.g. Archibald, 1993; Broselow & Park, 1995; Brown, 1993a, 1996, to appear b; Pater, 1993; Weinberger, 1994). Using the tools of current phonological theory, this new line of research has begun to define more precisely the role of the native language in the process of acquiring a second language; we are now in a position to develop a theory of L2 phonological interference which includes a principled explanation for the existence of L1 influence in some instances and its absence in others, as well as a description of the mechanism(s) by which this influence is exerted.

Building on the insights of this prior research, this chapter develops a model of speech perception, couched within current phonological theory, that accounts for the influence of the native grammar in both infant and adult speech perception. More specifically, by utilizing the theory of Feature Geometry, the proposed model demonstrates how the monotonic acquisition of phonological structure by young children restricts their sensitivity to particular non-native contrasts and how the continued operation of this existing phonological structure in adult speech perception constrains which non-native contrasts adult learners will be sensitive to in the L2 input and, therefore, capable of acquiring. By forging a link between infant speech perception and phonological acquisition, this research lays the foundation for a unified theoretical account of the interrelation between phonological acquisition and speech perception in children and adults. It is also offers an explanation for why learners perceive L2 sounds in terms of their native phonemic categories; by isolating and characterizing those phonological properties of the L1 that impinge upon L2 acquisition, this research identifies why and how this equivalence classification takes place. Finally, by demonstrating how the L1 grammar can both facilitate and hinder acquisition, these findings provide an answer to one of the questions currently central to second language acquisition theory: what determines partial L1 influence? The model outlined in this chapter accounts for the differential success that speakers of different L1's have in acquiring a given non-native contrast; it also accounts for the differential success that speakers with the same L1 have in acquiring various non-native contrasts. Furthermore, the experimental studies reported here demonstrate how the existing phonological system may block accurate perception of the input, thereby preventing the acquisition of novel segmental representations; it also establishes the circumstances in which the native grammar actually facilitates perception of non-native contrasts, demonstrating that when there is sufficient intake to the acquisition device, novel segmental representations can be successfully acquired.

We will begin by reviewing some of the previous research that has been conducted on the L2 acquisition of segments in order to set the context for the present research program and see why a new analysis is needed. Next, the relevant aspects of phonological theory will be laid out and explained. This will be followed by an examination of the development of the native phonological and perceptual systems, which will then lead us to a theory of phonological interference. After the implications of this theory for second language acquisition are laid out, the results of three experimental studies which test this theory will be reported and discussed. The paper concludes by considering some of the implications of these experimental data for the theory of phonological interference developed here as well as our theory of second language acquisition.

4.1 Historical Context & Theoretical Background

4.1.1 Previous research

Conducting research in applied areas such as acquisition requires one to strike a delicate balance between (at least) two continually developing theories: our theory of acquisition and our theory of grammar. In the case of L2 phonological acquisition, we must integrate insights from the theory of second language acquisition and current phonological theory. Advances in one of these usually requires us to reinterpret implications of the other in light of these new developments and to recast our theoretical models and experimental hypotheses. Similarly, failure of our acquisition models to correctly account for some aspect of the data force us to consider whether it is the acquisition theory or the linguistic theory underlying our model which needs to be modified. This complex bi-directional relationship often leads to a non-linear flow of progress in acquisition research. We are now once again at a point of reinterpretation, forced by the limitations of current models to reformulate our theory of L2 phoneme acquisition in terms of shifts within both the theory of segmental representation and the theory of second language acquisition.

Successful acquisition of phonological representations requires accurate perception of phonemic contrasts in the input; it is therefore clear that a comprehensive model of L2 phoneme acquisition must integrate not only a theory of second language acquisition and a theory of phonological representation, but also a theory of speech perception. Thus, it is not enough to ask only how the existing phonological system affects acquisition of L2 segments; we must consider all of the relationships in (1).

(1)

MATURE PHONOLOGICAL SYSTEM

↙ ↘

SPEECH PERCEPTION ⟷ ***PHONOLOGICAL ACQUISITION***

The majority of research on L2 phonological acquisition has investigated the relationship between the mature phonological system and phonological acquisition. But, the interrelation of these factors raises three additional issues that an adequate theory of L2 phoneme acquisition must explain: 1) how does the mature phonological system affect speech perception? 2) how does speech perception affect phonological acquisition? and (conversely) 3) how does phonological acquisition affect speech perception? By isolating the specific research questions addressed by previous L2 phoneme research and highlighting the particular theory of acquisition and/or theory of phonological representation assumed by each approach, we will see why a new analysis is needed.

The earliest systematic approach to the acquisition of L2 segments was undertaken within the contrastive analysis framework, the prevailing theory of second language acquisition of the time (Lado, 1957; Lehn & Slager, 1959; Stockwell & Bowen, 1965; Weinrich, 1953). The primary question addressed by this research was how the L1 influenced the acquisition of L2 segments, where acquisition was measured by the learner's ability to produce those segments. Guided by the premise that L1 structures heavily influence L2 learning, these researchers compared L1 and L2 phonemic inventories and hypothesized that L2 learners would have difficulty with those L2 sounds not found in the L1. Learners were assumed to substitute the "closest" L1 sound for a missing L2 sound in their productions, where "closeness" was based on the distance between the L2 sound and the L1 substitution on a language-independent segment chart. For example, Japanese learners of English, lacking the interdental fricative [θ] in their L1 inventory, would substitute their L1 phoneme [s] for it in production. This approach, however, was unable to account for aspects of the observed acquisition data. In particular, it incorrectly predicted that an L2 learner would have the same degree of difficulty with any and all of the L2 sounds not present in the L1 inventory, when, in fact, learners' performance on different L2 segments in experimental conditions ranges from native-like levels of accuracy to chance performance (see Munro, Flege & MacKay, 1996, for a detailed discussion of this point). Since this approach assumed a language-independent segment chart to predict substitutions, it also failed to explain why learners with different L1s would substitute different L1 sounds for a given L2 sound (e.g., Japanese speakers substitute [s] for [θ] but Russian speakers substitute [t], despite the fact that these L1s contain both /s/ and /t/, Hancin-Bhatt, 1994). These shortcomings, and in fact the most significant limitation of this approach, were due not to its comparison of L1 and L2 inventories, but rather to the level of phonological representation at which the languages where compared: these researchers took the phoneme to be the relevant unit of analysis.

Influenced by developments in generative phonology (and publication of Chomsky & Halle's *The Sound Pattern of English*, 1968), the next wave of research on L2 phoneme acquisition focused their analyses on the differences and similarities in distinctive features between the L1 and L2 (Michaels, 1973, 1974; Ritchie, 1968). Characteristic of these works was the assumption that sounds were made up of bundles of universal features and that these features could be ranked in a type of language-specific "feature prominence hierarchy" on the basis of how many phonemes each feature served to distinguish in that language's inventory. Although accurate production was still taken as the appropriate measure of acquisition, it was then recognized that perception might play a role in the transfer process. Specifically, language transfer was hypothesized to occur when the learner perceives the L1 and L2 constructs to be similar (Kellerman, 1977). In the context of features, this meant that the distinctive features of the L2 would be selectively perceived by the learner according to L1 perceptual biases. Thus, according to this line of research, difficulty with particular L2 sounds, in particular sound substitutions in production, could be explained in terms of featural differences between the L1 and L2, combined with the learner's perceptual biases. This line of research constituted an advance over the previous contrastive analysis approach in that its focus on the distinctive feature as the relevant unit for comparing the L1 and L2 provided language-internal evidence for differential substitutions. Moreover, it represented the first attempt to address the issue of how the mature phonological system might affect speech perception and how that, in turn, might affect phonological acquisition; however, it did not attempt to formally articulate this L1–L2 perceptual mapping, nor did these researchers provide any experimental evidence for their claims about how the native grammar influenced perception.

In the 1970s and 1980s, several perceptual studies conducted with native speakers and language learners provided the necessary experimental evidence, demonstrating that phonemes are indeed generally perceived in terms of the speaker's native categories (Abramson & Lisker, 1970; Miyawaki et al., 1975; Werker & Tees, 1984b; Williams, 1977). Since that time, three models have been proposed to explain how L2 sounds are mapped onto L1 sounds. The first model we will consider restricts itself to the relationship between the mature phonological system and speech perception; it does not address how phonological acquisition relates to these two factors. Best (1993, 1994) has developed the Perceptual Assimilation Model (PAM) to explain the role that a speaker's L1 phonological system plays in the perception of non-native sounds. According to this model, non-native sounds are assimilated to a listener's native categories on the basis of their respective articulatory similarities (more specifically, the spatial proximity of constriction location and active articulators); the degree to which a non-native contrast can be assimilated to native categories determines how well (if at all) a listener will be able to perceive that non-native contrast. While Best's proposal is based on, and supported by, experimental perceptual data, it lacks precise objective criteria for determining how non-native contrasts will be assimilated into native categories. Thus, although Best's model describes the role that a speaker's L1 phonological system plays in the perception of non-native sounds, it does not provide an explanation for why or precisely how this mapping occurs.

The Speech Learning Model (SLM), developed by Flege (1991, 1992, 1995), attempts to explain how speech perception affects phonological acquisition. According to this model, there are two kinds of sounds: "new" and "similar". New sounds are those that are not identified with any L1 sound, while similar sounds are those perceived to be the same as certain L1 sounds. Flege suggests that although the phonetic systems used in production and perception remain adaptive over the lifespan and reorganize in response to sounds in the L2 input, a process of "equivalence classification" hinders or prevents the establishment of new phonetic categories for similar sounds. However, this model does not include a theory–based proposal as to how L2 sounds are equated with L1 sounds. Thus, these two models attempt to elucidate the interrelation between the mature phonological system, speech perception and L2 phonological acquisition. However, despite their claim that there is an underlying mechanism that maps L2 sounds onto L1 categories, they fail to articulate the nature of that mechanism or adequately formalize the perceptual mapping process.

The most extensive model of speech perception–phonological acquisition interaction to be proposed thus far is Hancin-Bhatt's Feature Competition Model (FCM) (1994a,b). Expanding on the earlier work by Ritchie (1968) and Michaels (1973) described above, this model assumes that the features utilized in a grammar differ with respect to their "prominence": features (and feature patterns) used more frequently in the language's phonology will be more prominent than less frequently used features. Those features that are more prominent in the L1 system will tend to have a greater influence on learners' perception of new L2 sounds; that is, the feature prominences in the L1 will guide how L2 sounds are mapped onto existing L1 categories. Thus, like the PAM and SLM, the FCM assumes that L2 sounds are assimilated to L1 categories, yet this model goes one step further by providing an algorithm for determining feature prominence and, thereby, generating testable predictions for differential perception and substitution of interdentals across learners with different L1s. Furthermore, it is the first comprehensive model to investigate both the relationship between the mature phonological system and speech perception and the relationship between speech perception and the acquisition of L2 phonemic representations. Thus, this model addresses two of the three relationships indicated in (1) above; it also provides a more formal articulation of the L1–L2 perceptual mapping. However, to date its scope has been limited to the study of interdental substitutions; it is not clear whether this model can account for substitutions of other types of segments cross-linguistically or whether it can account for differential difficulty that speakers of a single L1 encounter in the acquisition of various L2 segments. Most importantly, the FCM does not address the reciprocal relation between perception and acquisition, namely how (L1) phonological acquisition affects speech perception. So, while we are moving closer and closer to a formalization of the influence that the mature phonological system has on speech perception (and the consequence of this for L2 acquisition), we still do not understand how the interrelation between speech perception and phonological knowledge originates; therefore, we fail to capture the essential nature of the phonological transfer mechanism. Investigating the development of speech perception and phonological acquisition in young children will enable us to explain why the mature phonological system exerts such a profound influence in adult speech perception; moreover, utilizing the tools of current phonological theory will allow us articulate the L1-L2 perceptual mapping mechanism more precisely.

4.1.2. Phonological theory & the representation of phonemes

Whereas previous research on L1 phonological interference primarily considered the phonemic categories of a language, phonological theory within the generative framework assumes that phonemes themselves have an internal structure. Thus, one way current phonological theory provides greater insight into the phenomenon of L1 influence is the distinction made between phonological representations and the components that comprise those representations. L2 phonological researchers now have an additional tool of analysis: the internal sub-components of phonemes constitute a further level of linguistic knowledge which may impinge upon L2 acquisition. However, these components (i.e., distinctive features) are not simply unordered bundles, as was assumed in the SPE framework (and theories of L2 phoneme acquisition couched within this framework). Instead, the distinctive features are themselves structured – an advancement which has implications for our theory of speech perception and phonological acquisition. Since an understanding of the internal structure of phonemes is necessary for the subsequent discussion of phonological interference, we will begin with a brief review of the relevant aspects of the theory of segmental representation assumed here.

Part of the phonological knowledge a speaker possesses about his or her language is what phonemes are contained in its inventory. However, as Hockett (1958:24) points out, a phonological system is "not so much a set of sounds as it is a network of differences between sounds." Thus, a native speaker's knowledge of the phonemes in his or her language is tantamount to knowing which segments are contrastive in that language. The most recent theory of segmental representation, known as Feature Geometry, produced significant insights into the internal structure of segments and, as a result, has advanced our understanding of segmental processes. In this framework the processes attested in natural languages are now believed to follow from the representation of the segment, in conjunction with a small set of principled phonological operations.

According to the theory of Feature Geometry, phonemes consist of distinctive features which are organized into a systematic hierarchy of constituents (Clements, 1985; Sagey, 1986).[2] Each phoneme has a unique structural representation (i.e., feature geometry) that distinguishes it from other segments in an inventory. Much of the work in phonological theory, as in linguistic theory in general, has been guided by the presumption that redundant information (defined as that information that can be predicted or easily supplied by derivation, e.g., syllable structure) is absent from underlying representations (Chomsky & Halle, 1968). Within Feature Geometry this principle is also extended to segmental representations. One such theory of underspecification is Minimally Contrastive Underspecification. According to this position, a segmental representation contains only the information needed to contrast it from all other segments in the system; any further specification will be provided by a system of phonetic implementation (Avery & Rice, 1989; cf. Archangeli, 1984, 1988, on Radical Underspecification). Thus, the precise representation of a segment will depend entirely upon which segments it contrasts with in the particular inventory. For example, in English, where the lateral approximant /l/ and central approximant /r/ are contrastive, /l/ may be represented as in (2a) while /r/ may be represented as in (2b), omitting irrelevant structure (Piggott, 1993; Brown, 1993b, 1995).[3]

---

[2] I will assume, along with a growing number of researchers, that features are monovalent and that it is the mere presence of a feature in the representation of a segment that designates the active involvement of its corresponding articulator; likewise, the absence of a feature entails that the corresponding articulator is not active for a given segment (e.g., Anderson & Ewen, 1987; Avery & Rice, 1989; van der Hulst, 1989). For example, the voiceless segment /t/ will simply not contain the feature [voice] in its representation (that alone ensures that the vocal cords are not active for this segment), whereas the phoneme /d/ will be specified for the feature [voice].

[3] The representations for /l/ and /r/ assumed here differ from standard representations. The phonological feature [lateral] is generally assumed to distinguish laterals from non-laterals, in this case /l/ from /r/. However, Brown (1993b, 1995) argues that [lateral] is not tenable as a phonological feature and that the contrast between /l/ and /r/ is best captured in terms of Place features (see reference for specific theoretical motivation and empirical evidence to support this claim; see also Piggott, 1993, and Spencer, 1984, for this view). Importantly, the representations for /l/ and /r/ given in (2) provide

(2) a. /l/

ROOT

SV PLACE

approximant

b. /r/

ROOT

SV PLACE

approximant Coronal

The fact that these segments have different representations reflects the fact that they are contrastive phonemes in the language; the presence of [coronal] in the representation of only one of them is sufficient to distinguish these segments in the grammar.

Conversely, when these two segments are not contrastive in a language, they will not have distinct representations. For example, in Japanese, [l] and [ɾ] are freely varying allophones of a single phoneme, so there will be only one underlying representation for these two surface segments. This is given in (3).

(3) /ɾ/

ROOT

SV PLACE

approximant

Despite the fact that this Japanese segment is realized as a coronal (like the English /r/), in accordance with our theory of underspecification, it is not specified for the feature [coronal] because it does not contrast with any other coronal approximants.[4] The phonetic realization of this segment as an [l] or an [ɾ] (which, unlike the English /r/, is a flap) varies freely (International Phonetic Association, 1979; Vance, 1987). Thus, in this way a speaker's knowledge of which sounds in his or her language are contrastive is represented by distinctive segmental representations.

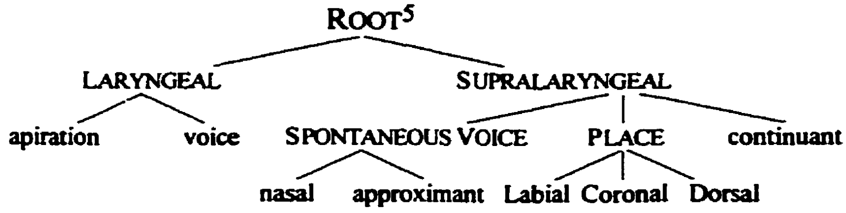
The full set of features manipulated in the world's languages and their dependency relations can be represented in terms of a single, universal Feature Geometry. This universal geometry is given in (4) for illustration.

---

an explanation for differential acquisition effects due to a speaker's first language grammar, which is not expected given the standard view of liquids.

4 The Japanese /ɾ/ will be distinguished from other coronal sounds (e.g., /s/) in terms of manner features.

(4)



This Feature Geometry is contained in the phonological component of Universal Grammar, the innate language faculty ascribed to the child by generative theorists. Like a syntactic principle or parameter, this geometry constrains the acquisition process and provides the learner with information about what phonemic oppositions are possible in natural languages. Thus, while no one language manipulates all components of this universal Feature Geometry, every phoneme in the world's languages can be represented in terms of the features and structural relations present in this geometry. Languages will differ with respect to their phoneme inventories and, hence, with respect to the set of phonological features they manipulate. However, the organization of those features, as given by the universal Feature Geometry, will be the same in every language. The learner's task is to determine which of the phonological features contained in this universal geometry are used to contrast phonemes in the language he or she is learning and to construct the appropriate representations. In the remainder of this paper, we will consider whether L2 learners can acquire non-native segmental representations as well as how, and to what extent, the native phonological system influences this process.

4.2 A Theory of Phonological Interference

In developing a theory of L1 influence, one of the issues we must address is *why* the L1 grammar exerts this influence. With respect to phonological interference, the relevant question is why foreign sounds are perceived in terms of the learner's native sound categories. In order to fully understand why the phonological system affects

---

[5] This model integrates properties of models proposed by Clements & Hume (1994), Piggott (1992) and Rice & Avery (1991b); however, the arguments and findings presented here do not hinge on the correctness of this particular hierarchical organization.

perception in this way, we must examine the ontogenetic development of these systems, as well as any interdependence between them. Studying the development of the L1 system will offer us insight into its operation in mature speakers; in addition, an understanding of how phonological knowledge is acquired in L1 acquisition will enable us to determine what conditions are necessary for successful L2 acquisition.

4.2.1. L1 phoneme acquisition

Since segments are distinguished in a grammar by their internal feature geometries, acquisition of a phonemic contrast involves the acquisition of the relevant structure (i.e. distinctive features) that differentiates those two phonemes (Rice & Avery, 1995, based on Jakobson, 1941). The child is able to contrast the two phonemes in his or her grammar once that phonological structure has been acquired and a representation has been constructed. Brown & Matthews (1993, in press) demonstrate experimentally that children's ability to differentiate phonemes phonologically develops gradually over time and in a systematic order that is consistent across children (see also Barton, 1980; Edwards, 1974; Garnica, 1973; Shvachkin, 1948, for related studies). Based on these results, they argue that UG provides the child's emerging grammar with a minimal amount of segmental structure (in fact, only those portions of the feature geometry that are universal) which is subsequently expanded over the course of acquisition until the adult feature geometry for the particular language is attained.

The systematic order of acquisition results from the hypothesized acquisition process and the nature of Feature Geometry itself; the child will only elaborate the feature geometry in his or her grammar in ways that are consistent with the hierarchical organization of features in UG. Specifically, the particular dependency and constituency relations that are encoded in the feature geometry in UG will be respected in the geometry posited by the child. For example, the presence of a dependent feature in a representation entails the presence of that feature's superordinate node. By extension, superordinate structure must be posited in the child's feature geometry before dependent structure can be elaborated. As a result, children will phonologically distinguish those segments that require less structure to differentiate before distinguishing those segments that require highly articulated structure. Thus, phonological structure is added to the child's grammar in a uniform, step-by-step fashion.

This step-by-step elaboration of the child's feature geometry in his or her grammar is driven by the child's detection of contrastive use of segments in the input (Jakobson, 1941; Rice & Avery, 1995). Once a child notices that two segments are used contrastively (i.e. are distinct phonemes), the phonological structure that differentiates the two segments is added to his or her grammar. If the child never perceives contrastive use of two segments (because, for example, they are allophones of a single phoneme in that language), the structure that differentiates them will never be posited. Therefore, the mere *presence* of two contrastive segments in the input (while necessary) is not sufficient to trigger acquisition; the learner must *detect* the contrast in the input.

4.2.2 Infant speech perception

In order for a learner to detect that two sounds are used contrastively, the learner must be able to discriminate the two sounds perceptually. Hence, proper development of the phonological system is dependent on properties of the speech perception mechanism. Given the fact that a child may be born into any language environment, it is imperative that he or she be equipped with adequate cognitive machinery to perceive the whole range of possible phonetic contrasts (cf. Burnham, 1986). Researchers have, in fact, demonstrated that infants as young as one month old are able to acoustically discriminate not only the sounds of the ambient language but many non-native contrasts as well (Eilers, Gavin & Oller, 1982; Eimas, Siqueland, Jusczyk & Vigorito, 1971; Streeter, 1976; Trehub, 1976; see Mehler, 1985, for a review).

Since the detection of contrasts in the input is crucial for the acquisition of phonemic representations, we need to consider how this capacity changes (if at all) as the child develops. In particular, whether or not an L2 learner has the capacity to perceive a non-native contrast will be a factor in determining if he or she will be able to construct the phonological representations necessary to distinguish the two segments phonologically. It is now well established that the ability to acoustically discriminate non-native contrasts decreases rapidly in infancy with exposure to a specific language, until the child is able to discriminate only those contrasts present in the language being acquired (see Werker & Polka, 1993, for a review of studies that establish this observation). In a series of studies, Janet Werker and her colleagues demonstrate that the decline in the ability to acoustically discriminate non-native contrasts occurs within the first year of life (Werker, Gilbert, Humphrey & Tees, 1981; Werker & LaLonde, 1988; Werker & Tees, 1984 a, b; and also Best & McRoberts, 1989; Best, 1994).[6] What is particularly fascinating is that this decline in perceptual capacity does not appear to be temporally uniform for all non-native contrasts. Experimental results indicate that perceptual sensitivity to certain non-native contrasts is lost before sensitivity to others, suggesting that loss of perceptual sensitivity to non-native contrasts is gradual and proceeds in a systematic order. An explanation for this decline in speech perception abilities, in particular one that integrates the role of linguistic experience, is still needed.[7] Both Werker and Best have tentatively suggested that the decline in the ability to discriminate some non-native contrasts may reflect the first stage of phonological development in the child, though neither is specific as to which aspect of the developing phonology might be responsible for this change. In the following section, I examine some findings from infant speech perception research and suggest a causal link between the development of a learner's feature geometry and the subsequent decline in perceptual capabilities. Establishing such a link will have important consequences for the acquisition of a second phonological system.

---

[6] Perception research has tended to focus on very young infants (0–14 months) or older children (4–12 years); there is a surprising lack of perceptual data for young children (1–3 years). Thus, while the decline in sensitivity to non-native contrasts has been shown to begin in the first year of life, it has not yet been determined whether this early perceptual reorganization is rigid or remains relatively flexible until the phonological system is firmly in place. For example, it has not yet been determined whether the observed early pattern of perception persists throughout language development or whether a child, if placed in the appropriate language environment once perceptual reorganization has begun, would regain the original sensitivities.

[7] See papers in Strange (1995) for reviews of the relevant speech perception data as well as several interesting proposals regarding the relationship between linguistic knowledge and the developing speech perception system.

4.2.3. The role of the L1 phonological system in speech perception

If we consider the findings from the infant speech perception research together with the research on phonological acquisition, an interesting parallelism emerges. We see that infants' perceptual capacities gradually *"degrade"* from all potential contrasts to only native contrasts (with some interesting exceptions), while their ability to discriminate segments phonologically gradually *improves* from no contrasts to only native contrasts. An exhaustive comparison of the stages of phonological and perceptual development is not feasible at this point, due to the limited number of non-native contrasts that have been investigated thus far. However, an examination of the data that we do have available suggests an intriguing possibility. According to Brown & Matthews, children first phonologically differentiate labials from velars, followed by labials from coronals. They do not distinguish segments that require a coronal node, such as /l/ and /r/, until relatively late. So, the node that distinguishes velar segments from other places of articulation is posited by the child before the node to distinguish among coronal segments is posited. Measuring auditory perception, Werker found the mirror order, with the perception of contrasts involving velars declining before contrasts involving coronals. Based on this convergence of the learner's perceptual and phonological capacities on the set of native sounds, Brown (1993, to appear b) proposes that there is a causal link between the learner's phonological development and the concomitant decline in his or her ability to acoustically discriminate non-native sounds.[8]

According to Brown's proposal, the acquisition of phonological structure (more specifically, the elaboration of feature geometry) in the child's grammar imposes upon his or her perceptual system the specific boundaries within which phonemic categories

---

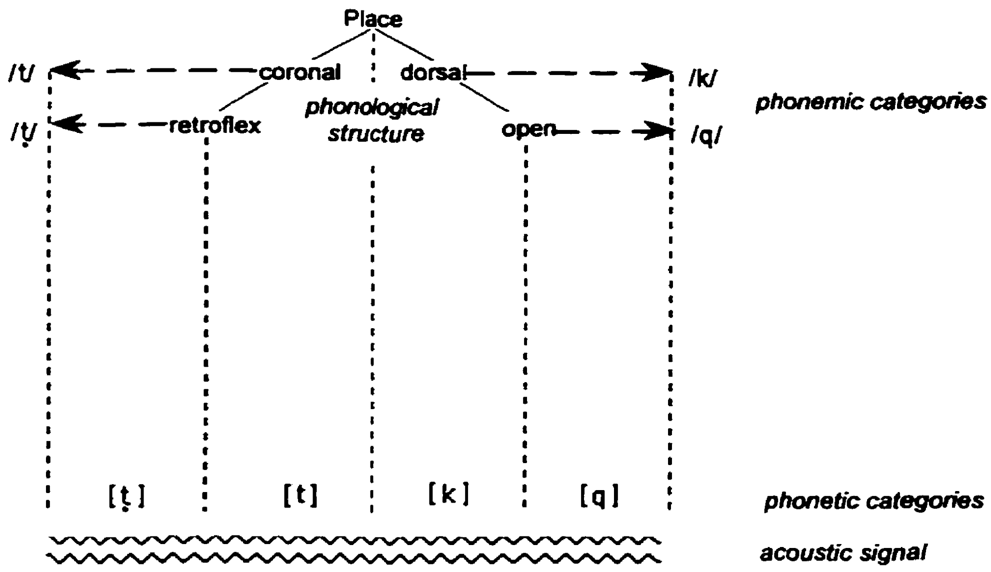
[8] There appears to be a time lag (approximately four to six months) between the age of the perceptual loss and the corresponding phonological development. A possible explanation for this time lag is that there is a confound between lexical development and phonological development, such that segmental representations are integrated into lexical items (which is what Brown & Matthews actually measured) shortly after they are first acquired. If this is indeed the case, acoustic discrimination tasks (specifically, lack of sensitivity) might provide a means of measuring the phonological development of children at even earlier ages than is currently available. This suggests that there may be an inventory of segments that is independent from the lexical items that contain them; this is a possibility that I will leave open for future research.

are perceived. In other words, the degradation of the perceptual capacities and the increase in the ability to distinguish sounds phonologically are the result of the same internal mechanism, namely the construction of phonological representations. This layer of phonological structure subsequently mediates between the acoustic signal and the linguistic processing system.

If we are correct in postulating that the acquisition of a phonological system determines the course of speech perception development, then it is reasonable to assume that the phonological system continues to constrain speech perception in adults. Mediating between the acoustic signal and the linguistic system, the phonological structure of the native grammar can be viewed as a filter which funnels acoustically distinct stimuli into a single phonemic category. This results in the well-documented phenomenon of categorical speech perception, whereby speakers of a language are able to easily distinguish members of different native phonemic categories and relatively unable to distinguish members of the same native phonemic category (Abramson & Lisker, 1970; Mattingly, Liberman, Syrdal & Halwes, 1971; Miller & Eimas, 1977; Pisoni, 1973; Repp, 1984). In other words, the mature speaker perceives the sounds of his or her native language, filtered through the existing phonological system, as distinct segments.

The hypothesis that the acoustic signal and phonological categories are mediated by a level of feature organization is quite intuitive given the fact that the acoustic signal cannot be characterized in terms of abstract categories, such as phoneme, but does correlate to properties of the gesture (e.g., place of articulation features correspond to spectral peaks in release bursts and to formant frequencies). The schematized diagram in Figure 1 (taken from Brown, 1993a, to appear b) illustrates the role of the intervening layer of phonological structure and how this level, in effect, funnels the acoustic signal into the phonemic categories of the speaker's language. The feature geometry depicted is from a hypothetical language in which /t/, /ʈ/, /q/ and /k/ are distinct phonemes; we will only consider the Coronal and Dorsal nodes (and their dependents) of the geometry.

Figure 1
Mediation of Speech Perception by Phonological Structure

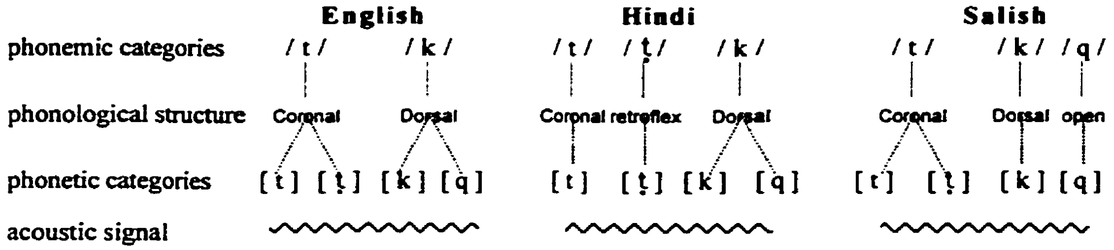


Starting from the bottom of the diagram, the acoustic signal is first broken down into phonetic categories. At this level, the acoustic signals for an alveolar [t] and a retroflex [ṭ] remain distinct. These stimuli then pass to the second level which consists of a speaker's feature geometry. This phonological structure serves to further categorize the phonetic stimuli into phonemic categories which are then fed into the language processor. Because this language exploits a dependent feature of the Coronal node, the phonetic signals for [t] and [ṭ] are channeled into the distinct phonemic categories /t/ and /ṭ/. The acoustic signals for [q] and [k] are processed in the same way. This model of speech perception is supported by research by Werker & Logan (1985), who found evidence for three distinct levels of processing: depending on the length of the interval placed between the stimuli (and hence on the memory load required to perform the task), subjects exhibit perception at either the auditory, phonetic or phonemic level. In particular, these researchers showed that, under certain conditions, English speakers are able to acoustically discriminate the Hindi /t–ṭ/ contrast more accurately than predicted by chance.

According to the model in Figure 1, the acoustic signal will first be divided into distinct phonetic categories, which are only subsequently categorized into native phonemic categories. Thus, regardless of the phonological system of a speaker, non-native contrasts are distinct at some level and may be discriminated under certain controlled conditions, as Werker & Logan have demonstrated. In other words, the "loss" of sensitivity observed in young infants is not really a loss at all, but rather is the result of perceptual reorganization – reorganization, I would like to suggest, that reflects the hierarchical organization of the feature geometry in the speaker's grammar.

The diagram in Figure 2 illustrates, in slightly more abstract terms, how the speech perception of English, Hindi and Interior Salish (Nthlakampx) speakers differs from one another.

Figure 2
Cross-language Speech Perception



In English, the Coronal node serves to distinguish coronals from non-coronals (e.g. /t/ vs. /p/), but no distinction is made within the coronal place of articulation (e.g. /t/ vs. /ṭ/). Thus, the English feature geometry does not contain the feature [retroflex]. As a result, (all) coronal sounds, regardless of their distinct acoustic signals, are perceived as a single phonemic category. Likewise, English makes no phonemic distinction between velar and uvular sounds; therefore, the Dorsal node has no dependents and velar and uvular sounds will be perceived as the English phoneme /k/. The feature geometry of Hindi also lacks the dorsal dependent [open], so that perception of [k] and [q] as /k/ is the same as for English speakers. Unlike English, however, the Hindi feature geometry contains both the Coronal node and its dependent [retroflex]. Thus, all coronal sounds will not be funneled into one phonemic category; the two features in the geometry ensure that /t/ and /ṭ/ will be perceived as distinct phonemes. In Interior Salish, the situation is just the reverse: /k/ and /q/ are perceived as distinct phonemes (due to the presence of the feature [open] as a dependent of the Dorsal node) while the acoustic signals for /t/ and /ṭ/ are perceived as the single phoneme /t/ (due to the absence of Coronal node dependents).

That the native system operates in this way is not accidental: perceiving speech in terms of phonemic categories undoubtedly aids processing and facilitates comprehension of the linguistic signal. Native speakers are continually faced with variable realizations of segments, due to coarticulation, sloppy articulation or interspeaker variability. By filtering out this irrelevant "noise" in the acoustic signal, the memory load put on the auditory system is greatly reduced and processing can proceed more quickly. Those variations in the acoustic signal that do not contribute to differences in meaning are simply not perceived by the listener. Yet, although categorical perception aids processing of one's native language, it can be a barrier to correctly perceiving and processing a foreign language: variation in the acoustic signal which is filtered out by the native phonological system (i.e., is treated as *intra*-category variation) may, in fact, contribute to differences in meaning in the foreign language (i.e., actually constitute *inter*-category variation). Thus, the influence of the mature phonological system on the perception of foreign sounds is an artifact of how speech perception functions in general. To summarize thus far, I have suggested that a learner's developing feature geometry causes the gradual decline in the ability to acoustically discriminate non-native contrasts and then continues to mediate between the acoustic signal and the linguistic processing system. The next section outlines the predictions this proposal makes for L2 acquisition of phonology.

4.2.4. Implications for L2 Phonological Acquisition

Establishing a link between a learner's phonological development and his or her speech perception has important implications for the acquisition of non-native contrasts by second language learners. In particular, this proposal suggests that the learner's native grammar constrains which non-native contrasts he or she will accurately perceive and, therefore, limits which non-native contrasts that learner will successfully acquire.

A speaker's phonological knowledge consists of phonemic representations as well as the features that comprise those representations. The position that the features exist in grammar (somewhere) independent of the segments they define is an assumption at this point in the discussion; however, we will see experimental support for this claim in Experiment 3 below. *A priori*, either of these levels of knowledge (i.e., featural or segmental) could potentially impinge upon the L2 acquisition process. According to the theory of phonological interference outlined here, however, it is the *features* contained in the learner's native grammar, not the phonological representations themselves, which constrain perception. The prediction of this position is that if a speaker's grammar lacks the feature that differentiates a given phonological contrast, then he or she will be unable to accurately perceive that contrast; conversely, the presence of the contrasting feature in the native grammar will facilitate perception of that non-native contrast, regardless of whether the particular segment is part of the inventory. That is, despite a lack of acoustic, phonetic or phonemic experience with a *particular* non-native contrast, a speaker's experience perceiving native phonemic contrasts along an acoustic dimension defined by a given underlying feature (for example, voicing) permits him or her to accurately discriminate *any* non-native contrast that differs along that same dimension. (This *is* a strong claim, but one I would like to maintain until empirical data force me to a weaker position). Thus, perception of certain non-native contrasts is possible by virtue of the fact that the phonological feature that underlies that particular acoustic dimension exists independently in the learner's native grammar.[9]

[9] Note that this is true regardless of the actual phonetic realization of a particular contrast. Take, for example, voicing contrasts: although languages may vary as to how they choose to acoustically realize the voicing contrast (i.e. actual VOT's may vary), since the same phonological feature underlies this contrast (i.e. [voice]), the claim is that speakers whose native language exploits this feature will be able to perceive all non-native voicing contrasts.

This position does not, however, entail that the phonological categories themselves play no role whatsoever in perception; while it is claimed that it is the features that determine the perceptual sensitivities (and that guide the mapping of the acoustic signal onto perceptual categories), it is still the existing phoneme categories which the incoming acoustic stimuli are sorted into, at least initially. As we will see from the experimental studies reported below, the effects of the L1 phonological categories will be most apparent in the initial stages of acquisition, before new L2 categories have been established. Nevertheless, it is the features of the L1 which ultimately enable or prevent the construction of these new L2 categories.

If the native phonological system affects perception of non-native contrasts in this way, either preventing or facilitating accurate perception, what are the consequences for the acquisition of these contrasts by L2 learners? Recall from our discussion of L1 phoneme acquisition, that acquisition of the relevant phonological structure is triggered by the learner's detection that the two sounds are used contrastively in the language (i.e. that they correspond to separate phonemes). For example, if the learner is to acquire the phonological structure required to differentiate /l/ and /r/ in his or her grammar, then he or she must notice that minimal pairs, such as *right* and *light*, are distinct words. In short, accurate perception of a phonemic contrast is necessary for successful acquisition of that contrast. It follows, then, that L2 learners will acquire only those non-native phonemic contrasts that they perceive as distinct sounds. If an L2 learner detects that two segments are used contrastively in the foreign language, then acquisition of the novel representations will be triggered.[10] On the other hand, if a contrast between two foreign sounds is not perceived (i.e. both sounds are perceived as belonging to the same phonological category), then acquisition will not be triggered and the L2 learner will fail to distinguish those segments in his or her

[10] The ability to construct novel segmental representations presumes, of course, that the acquisition device is still operative in L2 acquisition. See White (1989a) for arguments regarding the operation of Universal Grammar in L2 acquisition; see Brown (1994, to appear b) for a discussion of this issue with respect to L2 phonological acquisition.

interlanguage grammar. Put slightly differently, if the L2 input continues to be (inaccurately) mapped to L1 representations, there will be no impetus for acquisition.

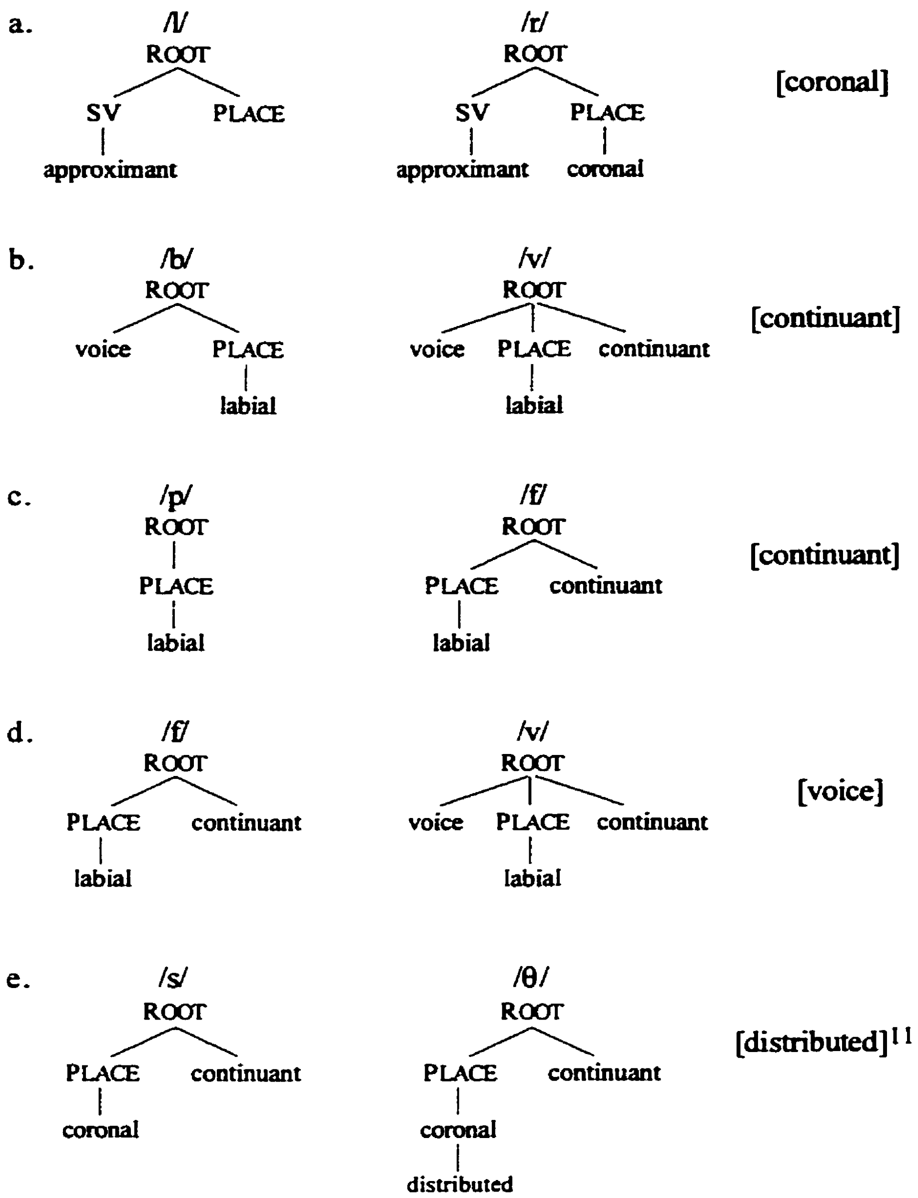
4.3 Experimental Evidence

The following three experimental studies investigate how the grammars of Japanese speakers, Korean speakers and Mandarin Chinese speakers affects their acquisition of English contrasts and whether, given the necessary conditions, novel segmental representations can be constructed. These studies were designed to explore three related issues: the acquisition of a range of phonemic contrasts by a single group of speakers (experiment 1), the acquisition of a particular contrast across different groups of speakers (experiment 2) and whether the nature of L1 phonological influence changes over the course of L2 development (experiment 3). As the studies examine the acquisition of different subsets of English contrasts, the representations of all of the segments under investigation, as well as the phonological properties of the three L1s, will be discussed together, prior to the description of the individual experiments.

4.3.1. Contrasts investigated

The English /l-r/, /b-v/, /p–f/, /f-v/ and /s–θ/ contrasts were chosen to test the proposed model of phonological interference because these pairs are not contrastive in Japanese, Korean or Chinese; furthermore, since these contrasts are distinguished by different phonological features, each contrast could potentially cause a differing degree of difficulty for these groups of learners (both with respect to the various contrasts and with respect to the various L1 groups). The internal structure of each pair is given in (5); note that these representations are for the segments as they occur in English. The phonological feature that distinguishes each contrast is given to the right (the superordinate SUPRALARYNGEAL and LARGYNGEAL components are not relevant for this discussion and have been omitted for ease of exposition).

(5) *Representations* *Contrasting feature*



These are the representations that the learner must acquire in order to distinguish these phonemes in his or her interlanguage grammar. The important thing to note is that each pair of phonemes is minimally differentiated by the presence of a single phonological

[11] The segments /s/ and /θ/ also differ acoustically in terms of stridency, and some phonologist distinguish them by the feature [strident]; however, following Kentowicz (1994:30), I assume that their phonological representations differ in terms of place features. The predictions for the learners of English will not differ, though, under either analysis, as neither [strident] nor [distributed] is an underlying feature in any of the three languages under investigation.

feature. Without this contrasting feature in the representation of one of the sounds, the two segments will not be distinguished in the learner's interlanguage grammar. The /l-r/ contrast is distinguished by the feature [coronal], the /b–v/ and /p–f/ contrasts by the feature [continuant], while the /f–v/ and /s–θ/ contrasts are distinguished by the features [voice] and [distributed], respectively. Whether an L2 learner will successfully acquire each of these contrasts depends entirely on the presence or absence of the contrasting feature in his or her native grammar.

4.3.2. Phonological properties of Japanese, Korean and Mandarin Chinese

Let us now examine the consonant phoneme inventories of Japanese, Korean and Chinese, given in (6), in order to ascertain whether the features that distinguish the English contrasts are contained in the mental grammar of these speakers (from Maddieson, 1984, and Vance, 1987).

(6) Japanese Inventory[12]

p	t		k	
b	d		g	
		tʃ		
		dʒ		
	s			h
	z			
m	n			
	ɾ			
w		j		

Mandarin Chinese Inventory[13]

p	t		k	
pʰ	tʰ		kʰ	
	ts	tʂ		
	s	ʂ		h
		ʐ		
m	n			
	l			
w	j			

---

[12] While other segments are realized phonetically in Japanese, such as [ɸ], they are derived (i.e., occurring in predictable phonological contexts) and do not, therefore, constitute independent phonemes.

[13] The claim that Mandarin Chinese does not contrast /l/ and /r/ phonemically requires some comment. This language contains /l/ and a segment which is transcribed in romanized script as "r"; this transcription gives the impression that there is a contrast between the lateral approximant /l/ and a central approximant /r/. This "r" segment, however, is classified by linguists as a voiced retroflex fricative, /ʐ/. For this study, I follow Maddieson (1984) in treating /ʐ/ as a voiced retroflex fricative and, crucially, not as a retroflex sonorant. But compare Rice (1992) who analyzes this segment as /r/ underlyingly; postulating that it surfaces as a voiceless retroflex fricative [ʐ] in onset position and as rhoticization of the vowel when in coda position [ɚ]. Thus, according to this analysis, /r/ and /l/ do contrast as sonorants in Mandarin. However, it is not clear that this analysis is correct. The coda position in Mandarin is restricted to nasals; thus it is unlikely that the rhoticization of the vowel is from the presence of an approximant in the coda position. Finally, only certain vowels are rhoticized (Chao, 1968; Wu, 1991). This suggests that the rhoticization is a property of the vowel itself, rather than the result of /r/ in the coda position.

Korean Inventory

p	t		k	
p^h	t^h		k^h	
p'	t'		k'	
		tʃ		
		tʃ'		
	s			h
	s'			
m	n			
	l			
w	j			

From this, we see that each of the five non-native contrasts we are interested in has a slightly different status *vis à vis* the inventories; the status of these contrasts is summarized in Table 1.

Table 1
Status of English Contrasts in Japanese, Chinese & Korean

	English Contrasts									
Status in Japanese Inventory	/b/ vs. /v/		/p/ vs. /f/		/f/ vs. /v/		/s/ vs. /θ/		/r/ vs. /l/	
is a native segment	√		√				√			
corresponds to a native segment				√[14]	√				√[15]	
does not correspond to a native segment		√				√		√		√
Status in Chinese Inventory										
is a native segment			√				√			√
corresponds to a native segment	√									
does not correspond to a native segment		√		√	√	√		√	√	
Status in Korean Inventory										
is a native segment			√				√			√
corresponds to a native segment	√									
does not correspond to a native segment		√		√	√	√		√	√[16]	

It should be noted that "corresponds to a native segment" means here that the surface realization of a given segment (in Japanese, Chinese or Korean) could reasonably be

14 Japanese contains a bilabial fricative [ɸ]; however, this is an allophone of /h/ and is realized before the high back unrounded vowel /ɯ/.

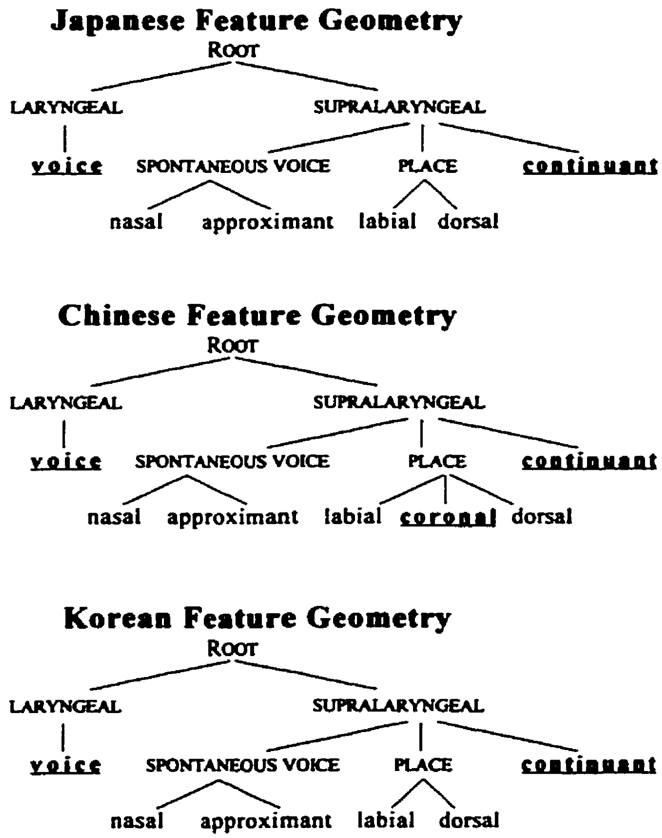
15 Japanese contains one liquid described as a flap [ɾ], which is not identical to the central approximant [r] in English, but is traditionally considered to correspond to English /r/, not /l/. This flap has several variants, which vary freely, one of which is phonetically similar to English [l] (Vance, 1987).

16 In Korean, [ɾ] and [l] are in complementary distribution, with [ɾ] (an apical flap) occurring intervocalically (Jung, 1962).

assumed to be a surface realization corresponding to the underlying representation of an English phoneme. For example, while Japanese does not contain a labiodental fricative (/f/), the Japanese bilabial fricative ([ɸ]) could correspond to the underlying representation of the labiodental fricative given in (5c). Thus, while Japanese, Chinese or Korean may not have a given English segment in its inventory, it may contain a very similar sound that could potentially factor into the acquisition of that English segment. Notice that, with respect to phoneme inventories, the status of these English contrasts is very similar for Chinese and Korean, which would lead us to expect to find comparable patterns of acquisition across these two groups of learners, if it is the phonemes of the L1 which constrain perception.

We can also use the consonant phoneme inventories in (6), along with our theory of underspecification, to determine which phonological features are used contrastively in these three languages; the adult feature geometries in (7) illustrate which features are manipulated.

(7)



The phonological features that we are interested in (i.e., the ones that differentiate the English contrasts at hand) appear bolded and underlined. Notice that the features [continuant], which distinguishes the /b–v/ and /p–f/ contrasts, and [voice], which distinguishes the /f-v/ contrast, are present in the grammar of all of three languages. Even though /b/, /v/, /f/ and /p/ are themselves not contrastive in these languages, other native segments differentiated by these particular features are contrastive. For example, the feature [continuant] is required in the Japanese grammar to differentiate native stop-continuant contrasts, such as the /t-s/ and /d–z/ contrasts, while the feature [voice] is present in the grammar in order to represent native voicing contrasts, such as /t-d/ or /s-z/. Thus, the feature that distinguishes /b–v/, /p–f/ and /f–v/ exists in the grammar for independent reasons.

However, the feature that distinguishes the /l-r/ contrast ([coronal]) is present only in the Chinese grammar (to distinguish the native alveolar /s/ and the retroflex /ṣ/); [coronal] is not present in the Japanese or Korean grammar as there are no consonants in either language that are distinguished from each other by this feature.[17] Finally, the feature that distinguishes the /s–θ/ contrast ([distributed]) is not utilized in any of the three grammars. Thus, in terms of features (as opposed to segments), Korean is more similar to Japanese than either is to Chinese.

4.3.3. Predictions

Recall that according to the theory of phonological interference being pursued here, it is the status of the contrasting *feature(s)* in the learner's native grammar (i.e. presence or

---

[17] Although there are coronal segments in Japanese (e.g. /t/, /s/, /n/), under a theory of Minimally Contrastive Specification, a feature will only be present in a grammar if that feature is required to contrast segments; accordingly, coronal segments in Japanese will be represented with a bare Place node. However, based on palatal prosody in Japanese mimetics, Mester & Itô (1989) argue that Japanese coronal segments are, in fact, represented with the feature [Coronal]. This specification is proposed in order to account for the fact that all coronal segments *except* /ɾ/ are palatalized (as are non-coronals). By specifying all coronal sounds, other than /ɾ/, with the feature [Coronal], the authors explain why /ɾ/ is not a target of this operation. The same facts, however, can be obtained by assuming (as the authors themselves do to explain why /ɾ/ cannot be geminated) that /ɾ/ is not specified for any Place Node at all, whereas coronal segments are specified for a bare Place Node (with no [Coronal] feature). Lacking a Place Node, /ɾ/ will never be the target of palatalization. This specification would also explain why coronals are the preferred target of this operation, with non-coronals becoming palatalized only in the absence of a coronal: since coronals lack Place features, the palatal morpheme has a free place to dock on these segments, whereas the addition of palatalization to the non-coronal segments creates a less-favored complex structure.

absence) that determines the perception and subsequent acquisition of non-native contrasts. Of particular interest is the difference between the acquisition of a non-native contrast when the second language learner's native grammar does NOT contain the feature that distinguishes the segments and the acquisition of a non-native contrast when the learner's native grammar DOES contain the distinguishing feature. Taking Japanese learners of English as an example, three of the contrasts under consideration, /b-v/, /p–f/ and /f-v/, are distinguished by a feature present in the Japanese grammar, whereas /l-r/ and /s–θ/ are not. Thus, we would expect the former three contrasts to pattern together in acquisition, to the exclusion of the latter two.

More specifically, since perception of a non-native contrast is facilitated by the presence of the relevant feature in the learner's grammar, Japanese speakers should accurately perceive that /b/, /v/, /p/ and /f/ are distinct segments. This is by virtue of the fact that the feature [continuant] operates in the mental grammar of Japanese speakers, functioning to sort acoustic stimuli that differ along this dimension. Likewise, Japanese speakers should accurately perceive /f/ and /v/ as distinct segments, in this case, because the feature [voice] exists in their grammar. Since accurate perception of these two non-native contrasts is facilitated by the learner's native grammar, the learner will detect that these sounds are contrastive in English and acquisition of phonological representations for these segments will be triggered. On the other hand, since perception of a non-native contrast is blocked by the absence of the relevant feature from the learner's grammar, Japanese speakers will be unable to accurately perceive a contrast between /l/ and /r/ or between /s/ and /θ/. Lacking the features [coronal] and [distributed], the phonological system of the Japanese speaker's grammar will funnel the distinct acoustic stimuli for /l/ and /r/ into one perceptual category and for /s/ and /θ/ into another. Consequently, Japanese speakers will perceive instances of /l/ and /r/ as the same sound (likewise for /s/ and /θ/). Unable to perceive that they are distinct segments, the learner will not detect contrastive use of these sounds and, as a result, novel representations will not be acquired; consequently, these segments will not be distinguished in the learner's interlanguage grammar.[18]

A summary of these predictions for the acquisition of the English contrasts by Japanese, Korean and Chinese learners is given in Table 2.

Table 2
Predictions for Acquisition of English Contrasts

	Japanese speakers		Chinese speakers		Korean speakers	
Contrasts	L1 contains feature	will perceive/ acquire	L1 contains feature	will perceive/ acquire	L1 contains feature	will perceive/ acquire
/b/ vs. /v/	continuant	YES	continuant	YES	continuant	YES
/p/ vs. /f/	continuant	YES	continuant	YES	continuant	YES
/f/ vs. /v/	voice	YES	voice	YES	voice	YES
/l/ vs. /r/		NO	coronal	YES		NO
/s/ vs. /θ/		NO		NO		NO

Three experimental studies that test these predictions are reported in the following section.

4.3.4. Experiment 1

Experiment 1 was designed to test whether the theory of phonological interference outlined in this paper could accurately account for variation in the acquisition of several different contrasts by learners with the same L1.[19]

Subjects

The experimental group consisted of fifteen Japanese speakers, ranging in age from 20 to 32 years, who had learned English as their only second language. Each of these

[18] While Japanese learners of English receive ample instruction in their language classes regarding the fact that /l/ and /r/ are contrastive in English, this type of explicit input, due to its very nature, does not feed into the acquisition device and, thus, does not trigger acquisition (Schwartz, 1993).

[19] As this study was originally reported in Brown (1993a, to appear b) a summary of the methodology and statistical analyses will be given; the reader is referred to the original study for more details.

subjects was raised in Japan and had come to North America to study in an undergraduate or graduate program at McGill University in Montreal, Canada, where the testing was conducted. The control group consisted of fifteen monolingual native speakers of English, who ranged in age from 15 years to 54 years. Their background information is summarized in Table 3.

Table 3
Subject Information

Group	Mean Age at Testing	Mean Age of Exposure	Mean Years Studied	Mean Years in N. America
Japanese	24.5	9	8	3.5
Controls	25	—	—	—

Contrasts investigated & hypotheses

The experimental contrasts were /l-r/, /b-v/ and /f-v/. A native contrast, /p-b/ was also included, as a control item for statistical comparison with the non-native contrasts. As outlined above, our hypothesis is that the Japanese speakers will accurately perceive the /b-v/ and /f-v/ contrasts, but not the /l-r/ contrast. Moreover, they should perceive these two non-native contrasts and the native /p-b/ contrast equally well. Again, successful acquisition of these two non-native contrasts, to the exclusion of /l-r/, is predicted to follow.

Tasks & materials

An AX Discrimination task was used to assess the subjects' ability to acoustically discriminate (i.e., perceive) the English contrasts. In this task, subjects hear a minimal pair (one item containing, for example, an /l/ and the other item containing an /r/ in onset position) and are asked to indicate whether the words are the same or different (e.g. *rip / lip*).[20] The items used in the test were natural tokens of real English

---

[20] Brown (1993a, to appear b) also examines acquisition of the /l-r/ contrast in onset clusters (e.g. glass/grass) and coda position (e.g. ball/bar).

monosyllabic words. These tokens were spoken by a man with a standard American English accent and taped so that the stimuli were identical for every subject.

Two types of foils were also included in the test materials. One foil type, which consisted of native contrasts, was included as a means of checking that poor performance on the task was not due to difficulty with the task itself; any difficulty with the task would be reflected in poor performance on the native contrasts as well as the non-native contrasts. A second type of foil, which consisted of identical pairs of words, was included to detect any response biases. Since each experimental minimal pair differs with respect to some consonant, the correct response for every trial is that the words are different. This set of foils ensured that any response bias or strategy toward responding that all of the pairs were different would result in inaccurate performance. If a subject responded that these identical stimuli were different, his or her data were discarded.

The second aspect investigated is whether those subjects who can acoustically discriminate the non-native contrast are able to acquire the phonological structure necessary to distinguish the two sounds phonologically. Based on the arguments in Brown & Matthews (in press), data from comprehension tasks are assumed to more be a more accurate indication of the learner's underlying competence than production abilities.[21] Subjects were, therefore, given a Forced Choice Picture Selection task (modified from Brown & Matthews, 1993, in press), in which the subject is presented

---

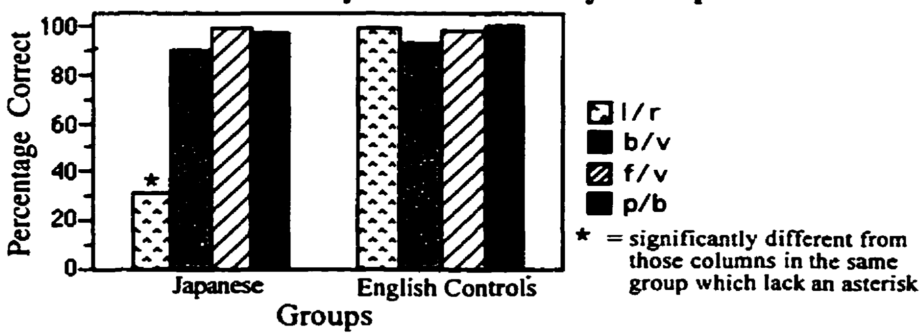
[21] Since production involves several peripheral mechanisms, such as motor control, relying on production data may lead us to underestimate or (particularly in the case of adults) overestimate the learner's underlying phonological competence. Several researchers have, in fact, demonstrated that L2 learners may be able to accurately produce a non-native contrast even though the same learners are unable to distinguish the two sounds perceptually (Briere, 1966; Flege, 1995; Goto, 1971; Sheldon & Strange, 1982). Japanese speakers, for example, have been shown to correctly articulate /l/ and /r/, despite their inability to perceive a difference between these two sounds. This is possible since adult learners have a developed motor control system and are able (with practice) to execute the necessary articulations. Once a speaker knows the spelling of a word that contains /l/ or /r/, he or she can accurately produce the correct liquid, thus giving the appearance of having acquired the contrast. To my knowledge, no one has investigated how this knowledge of proper articulation might be encoded into the learner's lexical representation of words. It is not clear whether this knowledge (which is dependent on orthography) is represented in terms of phonological structure. Thus, if we rely on production data we may falsely attribute more segmental structure to a learner's underlying phonological competence than he or she actually has. On the other hand, some L2 learners are more like young children in that they are unable to correctly produce a novel contrast despite their ability to perceive that contrast, in which case we would underestimate the learner's competence.

with two pictures and a verbal cue that corresponds to one of the pictures.[22] For example, the subject would see a picture of a *rake* on the left side of the page and a picture of a *lake* on the right side. At the same time, the subject would hear the word *lake*. The subject's task is to indicate which of the pictures the verbal cue names. In order to successfully complete this task, the learner must refer to his or her internal phonological representations of the pictured objects and determine which lexical representation corresponds to the verbal stimulus. If the subject's lexical representations of the pictured objects are identical (i.e. if they do not have the necessary phonological structure to contrast /l/ and /r/), then he or she will be unable to determine to which picture the verbal cue corresponds and should perform the task with chance accuracy. Successful completion of this task indicates that the subject has acquired the non-native contrast. The monosyllabic words used in this task were the same as those used in the AX Discrimination task. Both tasks were administered on the same day, with a short break between tasks.

Results & discussion

The graph in Figure 3 reports the mean performance scores of both groups on each of the contrasts.

Figure 3
Overall Auditory Performance by Group



[22] A training book was constructed which included every picture (one to a page) appearing in the experimental test. This book was used to familiarize the subjects with each of the pictures, and corresponding name, to appear in the Picture Selection task. This was done in order to minimize any errors that might be caused by the subjects' unfamiliarity with a particular stimulus item or illustration of an item. The materials used in stimuli preparation for the pictures were adapted from the Bilingual Aphasia Test (Paradis & Libben, 1987).

For the auditory discrimination task, a response that the two words in the minimal pair were "different" was counted as correct and a "same" response was counted as an error. Performance scores on each of the contrasts were tabulated separately for statistical analysis. From the Japanese subjects' near perfect performance on the native /p–b/ pairs, it is clear that the task itself does not pose any difficulty for the learners. Thus, performance on the non-native contrasts can be interpreted to reflect properties of the speakers' interlanguage grammar. As can readily be seen from the graph in Figure 3, the Japanese speakers were significantly poorer than the English controls at discriminating the /l-r/ contrast [t (28) = –16.16, p =.0001]. Yet, there was no statistical difference between the two groups in their ability to perceive the other contrasts; the Japanese speakers discriminated each of these English contrasts as accurately as the native controls [/b–v/ contrast: t (28) = –1.28, p = .21; /f–v/ contrast: t (28) = 1.87, p = .08; /p–b/ contrast: t (28) = –1.46, p = .15]. The Japanese speakers' performance on the /b-v/ and /f-v/ contrasts is quite striking: despite the fact that these are both non-native contrasts, they are very good (in fact, native-like) at discriminating each of them.[23] This suggests that these subjects perceived /b/, /v/ and /f/ as distinct speech sounds.

In order to evaluate performance on each contrast relative to the other contrasts, additional analyses were carried out separately on the two groups.[24] Beginning with the Japanese group, we find that their performance on the /l-r/ contrast is significantly worse than their performance on the other three contrasts [F (14, 45) = 119.85, p = .0001]; however, their performance on the other three contrasts (/b-v/, /f-v/, /p-b/) was not significantly different from one another. Thus, the Japanese speakers are unable to discriminate /l/ from /r/, perceiving them, instead, as a single category. In contrast,

[23] This result is perhaps even more surprising given the tendency of many Japanese learners to substitute /b/ for /v/ in production. But, as pointed out above, there is a well-known dissociation between comprehension and production skills, with comprehension assumed to be a more accurate reflection of the speaker's phonological knowledge.

[24] The subjects' individual data are further analyzed in Brown (1994) in terms of a standard binomial distribution, showing that the group data are indeed representative of each subject. This demonstrates that each of the Japanese subjects (not just the group as a whole) accurately discriminates the non-native /b-v/ and /f-v/ contrasts, but not the /l-r/ contrast.

though, not only do the Japanese speakers perceive the non-native /b-v/ and /f-v/ contrasts with native-like accuracy, as predicted, but they perceive them equally well. This is what we would expect, given that they are both distinguished by a feature in the Japanese grammar.[25] Moreover, that these non-native contrasts are discriminated as well as the native /p-b/ contrast suggests that perception of non-native sounds operates in the same manner as perception of native sounds. Although the native controls' performance on the /b–v/ contrast appears to be depressed relative to the other contrasts, there is, in fact, no statistical difference between the four contrasts [F (14, 45) = 1.44, p = .25]. Thus, we can regard the Japanese speakers' performance, in light of the native speaker data, as a true reflection of their perceptual capabilities.

To summarize, then, we have found a difference in the Japanese speakers' ability to perceive non-native contrasts, depending on whether the feature that distinguishes a given contrast exists in their grammar: the /l-r/ contrast (whose contrasting feature is absent from the Japanese grammar) is not accurately perceived, whereas the /b-v/ and /f-v/ contrasts (whose contrasting features are contained in the Japanese grammar) are accurately perceived, in a native-like manner (both with respect to the English controls and to the native Japanese contrast).

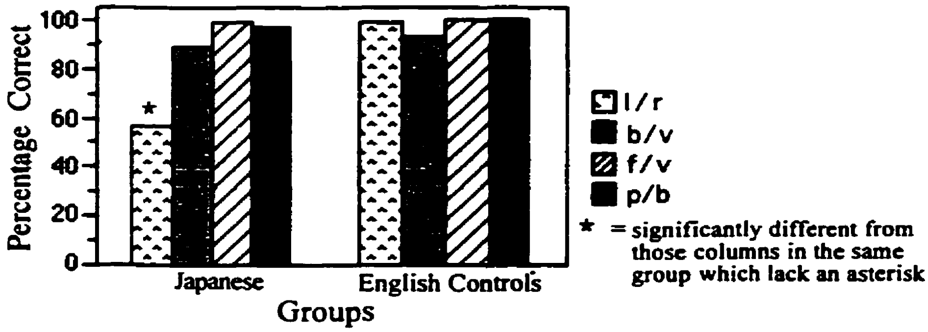
For the picture identification task, selection of the target picture was counted as a correct response, and selection of the contrast picture was counted as an error. The groups' overall performances are compared in Figure 4. Near perfect performance was attained by the Japanese group on the control items in this task. As in the auditory task, the Japanese speakers were significantly poorer than the English controls at differentiating the /l-r/ contrast [t (28) = –9.73, p =.0001]. There was, however, no statistical difference between the two groups in their ability to discriminate the other contrasts [/b–v/ contrast, t (28) = –1.8, p = .08; /f–v/ contrast, t (28) = –.32, p = .75; /p–b/ contrast, t (28) = –1.27, p = .22]. When shown two pictures that constituted a

---

25 Uniform performance is not predicted if the aspect of the native grammar responsible for filtering non-native sounds is the phonemic representations themselves (rather than the features). Recall from Table 1, that the members of the /b-v/ and /f-v/ contrasts have a different status *vis à vis* the Japanese phoneme inventory. Hence, if the phonemic representations constrain perception, we might expect differential perception of these two pairs, since the segment /b/, but not /f/, occurs in Japanese.

minimal pair (e.g., *rake*, *lake*), the Japanese subjects were unable to correctly choose the one that corresponded to the verbal cue. Yet these subjects performed this task with native-like accuracy when the pair of pictures differed by /b-v/ (e.g., *boat*, *vote*) or /f-v/ (e.g., *fan*, *van*).

Figure 4
Overall Picture Performance by Group



Let us now compare performance on each of the contrasts relative to each other. In this analysis, too, the Japanese speakers' performance on the /l–r/ contrast was significantly worse than performance on the other three contrasts, while performance on the /b-v/, /f–v/ and /p-b/ contrasts was uniform [$F(14, 45) = 57.65$, $p = .0001$]. This pattern of performance is confirmed by an examination of individual scores. Looking at the native controls, their mean performance on each of the contrasts was not significantly different from each other [$F(14, 45) = 1.56$, $p = .21$]. Thus, as with the auditory task, the performance of the Japanese subjects can be taken to accurately reflect their underlying phonological competence.

Although the performance of the Japanese subjects on the /l-r/ contrast was lower than their performance on the other contrasts, their accuracy rate (almost 60%) would seem to indicate that these learners have some knowledge of the /l-r/ contrast. However, in order to correctly interpret these results, it is necessary to consider the expected baseline performance on this task, that is, what chance performance would be. Suppose that a learner has no phonological knowledge of the /l-r/ contrast and is, therefore, unable to distinguish /l/ from /r/ in lexical representations. When that subject is presented with two pictures and a single verbal cue, he or she simply will be unable to decide which picture corresponds to the cue (i.e., since the representation for both items is the same, both correspond to the verbal cue). With a choice between two pictures, this subject has a 50% chance of choosing the correct one, just by guessing.[26] The observed performance, then, at 60%, is not significantly different from chance. We can infer with reasonable confidence from the Japanese speakers' performance on this task, that /l/ and /r/ are not differentiated in their grammars. Performance on the other contrasts, on the other hand, indicates that both the /b–v/ and /f–v/ contrasts are differentiated. In other words, the phonological structure that represents the /b-v/ and /f-v/ contrasts has successfully been acquired by these learners.

The hypothesis guiding this experiment was that perception of non-native contrasts is constrained by the phonological features manipulated in the native grammar of the learner. This led us to predict that Japanese learners of English would accurately perceive the /b-v/ and /f-v/ contrasts, as these two pairs are differentiated by features already present in the Japanese grammar, but that accurate perception of the /l-r/ contrast would be blocked by the absence of the relevant feature from the Japanese grammar. Each of these predictions was borne out by the data. The Japanese speakers' inability to perceive /l/ and /r/ as distinct phonemes can be understood as a direct consequence of the influence of the native grammar on the operation of the speech perception mechanism. The Japanese speakers' perception of /b/, /f/ and /v/ as distinct phonemes likewise provides experimental support for the model of phonological interference outlined in this paper. In a similar vein, since acquisition of a phonemic contrast is dependent upon accurate perception of that contrast, we predicted that Japanese learners would successfully acquire the /b-v/ and /f-v/ contrasts, but would fail to acquire the /l-r/ contrast. These predictions, too, were confirmed by the data. The Japanese learners successfully acquired only those non-native contrasts which they accurately perceived.

---

[26] Note that baseline performance on the AX Discrimination task is different from the picture task. According to the hypothesis that the Japanese grammar funnels the acoustic signal for both /l/ and /r/ into a single native phonemic category, Japanese speakers should perceive minimal pairs as identical. Thus, we would theoretically expect 0% accuracy at discriminating /l/ and /r/. In practice, though, they are able to correctly discriminate pairs more often – perhaps due to variations in duration and amplitude, which where not controlled for in this study. The difference in the subjects' performance on the two tasks (30% vs. 60%), then, is not indicative of differing abilities to perform each task, but rather reflects the fact that the baseline performance is different for each task.

The finding that Japanese speakers do not accurately perceive the difference between /l/ and /r/ is not particularly surprising, given the large body of literature that reports this observation (Goto, 1971; Miyawaki et al., 1975; Sheldon & Strange, 1982; Strange & Dittmann, 1984; Yamada, 1995). However this current research is the first to also examine and compare Japanese speakers' perception of additional English contrasts; it is also the first to investigate these speakers' acquisition of the /l/ and /r/ feature geometric representations.

The theory of phonological interference that has been tested in experiment 1 has correctly accounted for differences in Japanese learners' abilities to acquire different English phonemic contrasts. The theory further predicts that speakers of different L1s which differ in the features that they utilize will exhibit differing success rates of acquiring various contrasts. The following experiment tests whether this model of phonological interference can account for *cross-language* differences in L2 phonological acquisition.

4.3.5. Experiment 2

The purpose of this experiment was to examine differences in the acquisition of English contrasts by speakers of different native languages, and to replicate the findings in experiment 1 for Japanese speakers.

Subjects

A total of fifty-one subjects, divided into one control group and three experimental groups, participated in this study. One experimental group consisted of fifteen undergraduate Japanese speakers who were learning English at Hokkaido University, Sapporo, Japan and had never lived in an English-speaking country. The second experimental group consisted of fifteen native Mandarin Chinese speakers who were enrolled in graduate programs at Hokkaido University (and, therefore, proficient in Japanese). Eleven native speakers of Korean (also proficient in Japanese) comprised the final experimental group; these subjects were also graduate students at Hokkaido

University. Neither the Chinese native speakers nor Korean native speakers were enrolled in English classes at the time of testing. The control group consisted of ten native monolingual speakers of American and British English, who teach English at universities in Sapporo, Japan. Table 4 summarizes the background information for each of the four groups.

Table 4
Subject Information

Group	Mean Age at Testing	Mean Age of Exposure	Mean Years Studied
Japanese	20.3	11.8	8
Chinese	30.7	12.6	10.4
Korean	30	12.8	9.9
Controls	34.3	—	—

Contrasts investigated & hypotheses

Acquisition of the following contrasts was investigated: /p–f/, /f–v/, /s–θ/ and /l–r/, with the /p–t/ contrast (a native contrast for all groups) serving as a control item. Speakers of Chinese, Korean and Japanese were chosen for comparison because the grammars of these languages differ in interesting, theoretically-relevant ways. In particular, when we compare the status of English contrasts in L1 phoneme inventories (Table 1), Chinese and Korean appear to be more similar to one another, which might lead us to expect that speakers of these two languages would have the same difficulty (or success) acquiring the English contrasts under investigation. However, when the features employed in each of these grammars are compared (Table 2), Korean and Japanese are more similar to one another. Thus, examining acquisition by all three groups should provide evidence as to which level of phonological knowledge is responsible for L1 interference. To briefly review the predictions set out in Table 2, according to the theory of phonological interference adopted in this study, speakers of all three languages should accurately perceive and have successfully acquired the /p–f/ and /f–v/ contrasts since each of the L1 grammars utilizes the contrasting features ([continuant] and [voice], respectively) to distinguish native segments. Likewise, as the three L1 grammars lack the feature that contrasts /s/ and /θ/ ([distributed]), we predict that these two segments will be perceived (inaccurately) as a single category by speakers of all three languages; unable to hear a contrast between the two segments, they will also fail to acquire the feature geometric structure necessary to distinguish them phonologically in their interlanguage grammars. Finally, the three language groups should differ with respect to their ability to perceive and acquire the /l–r/ contrast: Chinese speakers, whose L1 contains the feature [coronal], will accurately perceive and, therefore, acquire this contrast; whereas, the adult feature geometry of Japanese speakers and Korean speakers will fail to sort the acoustic signal for these two sounds into distinct perceptual categories and their acquisition of the novel segmental representations will be prevented.

Tasks & materials

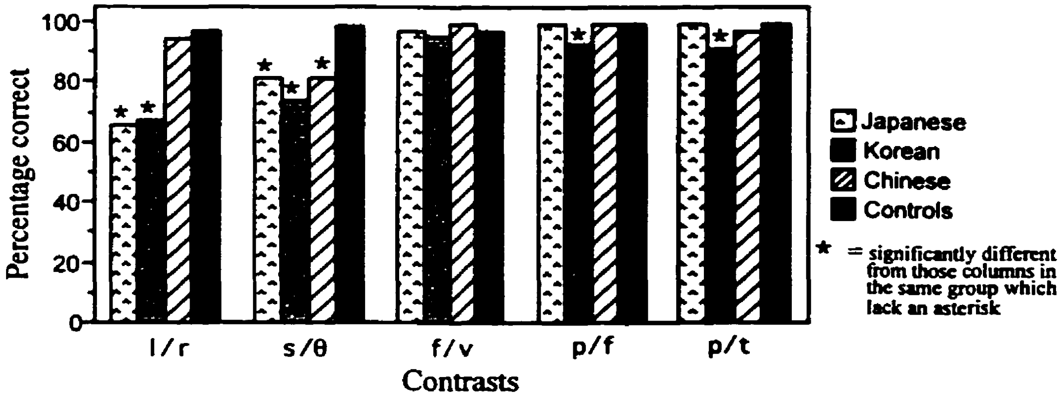
Phonological competence was assessed with the same Forced Choice Selection task used in experiment 1 above. A 4IAX discrimination task, rather than an AX task, was used to assess perception in this study. In the 4IAX task, each trial consists of two pairs of words (Pisoni, 1971); in one of those pairs, the two words will be different (i.e., a minimal pair), and the other pair of words will be the same (e.g., *ra/ra, ra/la*) . The subject's task is to indicate which of the two pairs of words is different. This task is becoming increasingly employed in speech perception research, since the AX task has been argued to bias the subject to respond "same" when discrimination is difficult (Beddor & Gottfried, 1995). The 4IAX task avoids this response bias since the subject knows that one of the pairs is, in fact, different and must simply determine which one.

The stimuli for this task, again in contrast to experiment 1, were non-words in order to prevent the subjects' perception from being influenced by their familiarity with particular lexical items (Yamada, Kobayashi & Tohkura, in press). The "same" pairs consisted of two instances of a CV syllable whose onset consonants were members of the same phonemic and phonetic category (e.g., aspirated [p^ha]), but which were not physically identical. Thus, subjects could not accurately choose the "same" pair (and thereby determine the "different" pair) simply by comparing physical objects and attending to non-linguistic acoustic variations, such as amplitude or speed. The "different" pairs consisted of two CV syllables whose onset consonants were members of different phonemic categories. Given these type of stimuli, accurate performance on this task requires the learner to filter out irrelevant variations across the segments and respond to higher-order phonological information. Stimuli were recorded by a male speaker of standard American English onto a Sony DAT Workstation and then arranged temporally by computer to create uniform intervals of 1000 milliseconds between members of a pair, 1800 millisecond intervals between pairs in a trial and 3000 millisecond intervals between trials.[27] These time intervals were chosen following Werker & Logan (1985) to ensure phonemic processing of the stimuli.

Results & discussion

For each 4IAX trial, selection of the pair whose members were from different phonemic categories was counted as correct and selection of the pair whose members were from the same phonemic category was counted as an error. The mean performance scores of all groups on each of the contrasts are reported in Figure 5.

Figure 5
Overall Auditory Performance by Contrast



[27] A second set of stimuli were recorded using two male speakers in order to compare one– vs. two–talker conditions. An identification task was also included. Results from the two–talker condition and the identification task will not be discussed here but can be found in Brown (to appear c), along with further methodological details.

The comparison that we are primarily interested in here is between the performance of the three language groups on particular/individual contrasts. As can be seen from the graph, the Japanese and Korean speakers are not as good as the Chinese speakers at discriminating the /l-r/ contrast. Statistical analyses reveal two distinct perceptual patterns: the Chinese speakers' performance is not significantly different from the native controls' performance, while the performance of the Japanese and Korean speakers is significantly worse than the Chinese speakers and native controls, but are not significantly different from each other [F (3, 47) = 16.39, p = .0001]. As we saw in experiment 1, the Japanese speakers are unable to distinguish spoken tokens of /l/ and /r/; we see that Korean speakers, too, perceive these segments as a single category (confirming Borden et al.'s, 1983, findings), whereas the Chinese speakers have no problem discriminating this contrast.

Turning to the /s-θ/ contrast, we find that the Japanese, Chinese and Korean speakers all discriminated this contrast equally poorly; they are significantly worse than the native controls, but not significantly different from each other [F (3, 47) = 3.8, p = .016]. The acoustic signals for these two sounds are funneled into the same perceptual category by speakers of all three language groups. With respect to the /f–v/ contrast, we also find consistent performance but, in this case, the groups are equally good, and not significantly different from the native controls [F (3, 47) = 1.49, p = .23]. The feature [voice] in the L1 grammar serves to separate, and keep distinct, the acoustic signals for these two sounds as they are processed. Performance on /p–f/, the other non-native contrast that is distinguished by an L1 feature manipulated by all three L1s, is roughly uniform for all groups, although we do find a small statistical difference between the experimental groups [F (3, 47) = 2.93, p = .04].

The Japanese and Chinese speakers are able to discriminate these two sounds as accurately as the native controls; however, the Korean speakers are significantly worse than both experimental groups and the native controls. However, the Korean speakers' performance, at 90% accuracy, is still well above chance and can be considered accurate, albeit not native like. Finally, all of the groups are able to accurately discriminate the native /p–t/ contrast, though small differences among groups do approach statistical significance [$F(3, 47) = 2.6$, $p = .06$]. Once again, the performance of the Korean speakers (91%) is a bit lower than the other groups. Given that /p–t/ is a native contrast for these speakers, this result is somewhat surprising and suggests that the Korean speakers' performance on all of the contrasts in this task is slightly depressed. Overall, the performance of the Japanese, Korean and Chinese speakers on the /f–v/ and /p–f/ contrasts is quite remarkable: despite the fact all three languages lack these two contrasts, these learners perceive them with native-like accuracy. Their ability to perceive these contrasts is particularly striking in light of their inability to discriminate the /s–θ/ contrast. Moreover, the differing ability of the Chinese speakers, on the one hand, and Japanese and Korean speakers, on the other, to accurately perceive the /l–r/ contrast indicates that the ability (or inability) to perceive non-native contrasts is linked to phonological properties of those contrasts and the L1 grammars, not to acoustic properties of the sounds themselves.

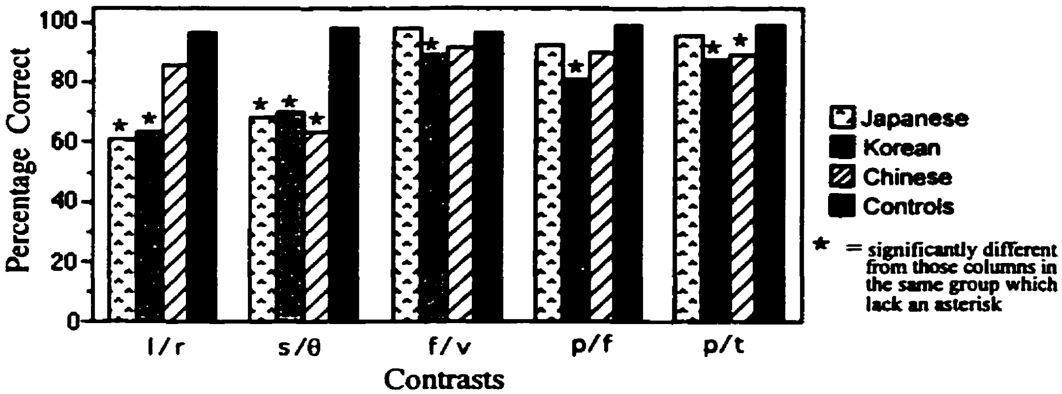
Although our main interest in this study is in differences between groups, it is still informative to consider performance on each of the contrasts relative to the others. In order to make such comparisons, additional statistical analyses were carried out separately for each group. It should be kept in mind that baseline performance on the 4IAX task is different than on the AX task. Recall that in the AX task the subject's decision is whether the two words are the same or different. If a subject cannot hear a difference between two sounds, then he or she will respond "same". In this case, performance would theoretically be 0% accuracy. In other words, given the influence of the native grammar, the probability of responding "same" or responding "different" is not 50%. In contrast, the subject's decision in the 4IAX task is which pair of sounds is different. If the subject's L1 grammar causes him or her to hear both pairs of words as being the same, the choice is still between "first pair" and "second pair" and the probability of randomly choosing either one is 50%. Thus, an inability to perceive a contrast in the AX task would result in 0% accuracy, whereas an inability to perceive a contrast in the 4IAX task would result in 50% accuracy. A consequence of this is that comparing performance on the 4IAX task across different phonemic contrasts is more difficult since differences between performance on contrasts that are perceived accurately and those that are not will be smaller (50% – 100%; cf. AX task: 0%–100%); for the same reason, scores from the two tasks cannot be directly compared.

Starting with the Chinese group, we find they are equally good at discriminating the /p–f/, /f–v/ and /l–r/ contrasts, and with the same accuracy with which they distinguish their native /p–t/ contrast; they discriminate all of these contrasts significantly better than they do the /s–θ/ contrast [$F(14, 60) = 8.55$, $p = .0001$]. This is what we would expect, given that the former non-native contrasts are distinguished by a feature contained in the Chinese grammar, whereas the latter is not. The Japanese group, too, discriminate the /p–f/ and /f–v/ contrasts with the same accuracy that they discriminate their native contrast, and they are significantly better at perceiving these contrasts than they are the /s–θ/ or /l–r/ contrasts [$F(14, 60) = 29.78$, $p. = .0001$]. These speakers do not, however, perceive the /s–θ/ and /l–r/ contrasts equally poorly; their discrimination of /l–r/ is worse than their discrimination of /s–θ/. It appears that, even though the acoustic signals for both sets of sounds will each be funneled into their respective category and perceived as the same sound, in a temporally adjacent presentation, the Japanese speakers are able to distinguish /s/ and /θ/ (but not /l/ and /r/) with higher accuracy than would be predicted by chance (possible reasons for this difference are discussed below). We find a similar pattern with the Korean speakers: /l/ and /r/ are discriminated less accurately than /s/ and /θ/ and performance on both of these contrasts is significantly worse than on the other contrasts [$F(10, 44) = 9.06$, $p. = .0001$].

The most important thing to note from these perceptual data is that Japanese speakers and Korean speakers differ from Chinese speakers in their ability to discriminate /l/ and /r/. This difference between the language groups might seem surprising given that all three languages lack this phonemic contrast. However, it can be properly understood as a direct consequence of the influence of the phonological features in their respective native grammars: the presence of the feature [coronal] in the grammar of Chinese speakers ensures that acoustic stimuli which differ on this dimension will be perceived as distinct, whereas the absence of the feature from the Japanese and Korean grammars causes the acoustic signal for these two sounds to be funneled into a single perceptual category. The three language groups do not differ in their ability to accurately discriminate those contrasts which are distinguished by a feature that exists in all three L1s or in their inability to perceive those contrasts which are distinguished by a feature not utilized in their native grammars.

Figure 6 compares the groups' overall performance on the picture selection task. With respect to the /l-r/ contrast, we find that, as in the auditory task, the Chinese speakers perform more accurately than the Korean speakers and the Japanese speakers [$F(3,47) = 21.35$, $p = .0001$]; in fact, they perform as well as the native controls. Chinese speakers have no problem choosing between two pictures that constitute a minimal /l-r/ pair, indicating that these two phonemes have distinct representations in their interlanguage grammars. The Japanese and Korean speakers, however, were significantly worse on this contrast than the Chinese speakers and native controls, though not different from each other; thus, neither of these two groups of speakers distinguishes /l/ and /r/ phonologically.

Figure 6
Overall Picture Performance by Contrast



All three groups of learners were unable to perform this task accurately when the lexical items differed by /s-θ/; there was no difference between experimental groups, and their performance was significantly lower than the controls' performance

[F (3, 47) = 11.53, p = .0001]. This indicates that a new segmental representation for /θ/ has not been acquired by the learners, as we predicted, so /s/ and /θ/ are represented by the same geometric structure in their interlanguage grammars. In contrast, lexical items containing /f/ and /v/ are distinguished phonologically by all three groups of learners, as indicated by their high performance levels, though the Korean speakers' performance, at 89% accuracy, is slightly worse than the native controls' performance [F (3, 47) = 3.21, p. = 03]. Similarly, on the /p–f/ contrast, the Japanese and Chinese speakers are as accurate as the controls, while the performance of the Korean speakers, though significantly lower, is still well above chance (83%) [F (3, 47) = 5.9, p. = .002]. With respect to the native /p–t/ contrast, all language groups distinguish /p/ and /t/ in their interlanguage grammars. Japanese speakers distinguished these two sounds in a native-like fashion; however, both the Chinese and Korean speakers were just slightly less accurate than the native speakers [F (3, 47) = 3.22, p. = .03].

Overall, then, we see that the learners in all three groups have distinct segmental representations for /p/, /f/ and /v/, while the Chinese speakers also have distinct representations for /l/ and /r/, and none of the learners have distinct representations for /s/ and /θ/. Let's now compare performance on the different contrasts by each group individually to confirm these acquisition patterns.

A separate analysis of the Chinese group reveals that the /s–θ/ contrast is distinguished much more poorly than the other contrasts, including the native /p–t/ contrast [F (14, 60) = 13.4, p = .0001]; the /l–r/, /p–f/ and /f–v/ contrasts, however, are distinguished equally well and as accurately as the native /p–t/ contrast. This means that /l/, /r/, /p/, /f/ and /v/ each have a distinct segmental representation in the Chinese speakers' interlanguage grammars; /s/ and /θ/, on the other hand, will correspond to the same phonological structure and, therefore, will not be distinguished in these learners' interlanguage grammars.

Analysis of the Japanese data also confirm two distinct acquisition patterns [F (14, 60) = 34.99, p. = .0001]. These learners represent the /p–f/ and /f–v/ contrasts in their interlanguage grammars in the same way that they represent the native /p–t/

contrast; there is no statistical difference between their (equally good) performance on these three types of contrasts. There is also no difference in their (in)ability to phonologically distinguish the /l-r/ and /s–θ/ contrasts: they are equally poor. This finding is especially interesting, given the difference we found between these two contrasts on the auditory discrimination task. Despite the slight advantage in perceiving /s/ and /θ/, it is not sufficient to trigger acquisition. Neither the segmental representation for /l/ nor for /θ/ has been acquired by these learners.

Finally, we turn to the Korean group, who have the identical pattern of acquisition as the Japanese speakers: the /l–r/ and /s–θ/ contrasts are not distinguished in lexical items, whereas the /p–f/ and /f–v/ contrasts are [F (10, 44) = 12.8, p. 0001]. In fact, although we saw above that the performance of the Korean speakers was slightly depressed on the /p–f/ contrast relative to the other language groups, it is not significantly different from their performance on their native contrast. Thus, it appears that whatever is causing the lower performance on the non-native contrast is not due to the non-native nature of the contrast, but rather to some more general performance factor. Nevertheless, additional studies examining Korean speakers' auditory and phonological discrimination are clearly required to establish their perceptual and linguistic abilities conclusively.

In summary, this experiment was conducted in order to determine whether our theory of phonological interference could account for the acquisition of English phonemes by speakers of different languages. Assuming that perception of non-native contrasts is constrained by the phonological features manipulated in the learner's native grammar and that languages differ as to the features they manipulate, we would expect learners with different L1s to differ in their ability to acquire particular non-native contrasts. Japanese, Korean and Chinese differ in just this way.[28] Specifically, the

---

[28] The differential performance of the two language groups also speaks to phonological theory, providing experimental evidence for the representations of /l/ and /r/ assumed in this paper. According to Brown (1993b, 1995), /l/ and /r/ are distinguished by the presence of Coronal in the representation of /r/. We can interpret the differential performance of the two groups in terms of the presence of this feature in Chinese and the lack of it in Japanese. However, according to Rice & Avery (1991b), /l/ and /r/ are differentiated not in terms of place features, but by manner features: /r/ contains a vocalic node

grammar of Chinese contains the feature [coronal], whereas the grammars of Japanese and Korean lack this feature. Given this, speakers of Japanese and Korean, on the one hand, and Chinese, on the other, should differ in their ability to acquire the /l-r/ contrast, which relies on the feature [coronal]. This is, indeed, what we found. Chinese speakers accurately perceive this contrast and, therefore, successfully acquire it. Japanese and Korean speakers are unable to acquire this contrast since they do not perceive /l/ and /r/ as different segments.

These three groups were fairly evenly matched for age of exposure, education and years spent studying English. Therefore, we can be confident that the differential performance of these groups stems from their respective L1s. Simply comparing the phoneme inventories of these three languages, however, does not allow us to explain why Chinese speakers accurately perceive and acquire the /l–r/ contrast, but that both the Japanese and Korean speakers do not. By the same token, it is only by considering the features utilized by the L1s that we can adequately explain why all three language groups were able to perceive and acquire the /p–f/ and /f–v/ contrasts, which are distinguished by features that exist in Japanese, Chinese and Korean.

Likewise, the absence of the relevant feature from all three L1s accounts for their uniform inability to acquire the /s–θ/ contrast. Thus, the differential abilities of the Japanese and Chinese speakers lends support to our theory of phonological interference: not only can we account for disparate acquisition of non-native contrasts by speakers of a single language, we can also explain disparate acquisition of a particular non-native contrast by speakers of different languages.[29]

Now that we have seen how the native grammar can affect perception and acquisition of non-native contrasts, a question that naturally arises is whether the native grammar always constrains phonological acquisition in this way or whether its effect changes over time, as the learner progresses. The Japanese learners tested in experiment

---

whereas /l/ does not. Chinese and Japanese do not differ with respect to this feature, thus Rice & Avery's representations incorrectly predict that Chinese and Japanese speakers should perform similarly.

[29] These conclusions are supported by Brown (to appear b) which compares the auditory and phonological discrimination abilities of Japanese and Chinese speakers living in North America and Brown (1996) which compares Japanese and Chinese speakers living in Japan.

1 were relatively advanced, living in North America and receiving abundant natural English input. Since these learners had already acquired two of the three non-native contrasts, we found no evidence for any stages of acquisition. The learners tested in experiment 2 had never lived in an English-speaking environment and were receiving minimal to no aural English input at the time of testing; but, although these learners were not always as accurate as the native speaker controls on those contrasts they had acquired, we still did not observe distinct stages of acquisition. Since the first two experiments were not longitudinal and also did not compare learners with differing levels of L2 proficiency, we have no data to determine whether there is any change in learners' perceptual capacities. Does perception of non-native contrasts improve over time? Is there any effect of increased linguistic input? Is there evidence for stages of acquisition? These questions were addressed in the following experiment, which investigated the acquisition of English contrasts by low proficiency and higher proficiency Japanese learners of English.

4.3.6. Experiment 3

This experiment was conducted in order to determine whether the influence of the native grammar on the perception of non-native contrasts changes over time as the L2 learner progresses.[30]

Subjects

The subjects for this experiment were thirty-five native speakers of Japanese and 10 native speakers of English. The control group comprised American, British and Canadian English teachers at Hokkaido University and Hokkai Gakuen University, in Sapporo, Japan. The Japanese subjects were learning English as a foreign language at Hokkaido University and had never lived in an English-speaking country. Based on teacher interview assessment of their overall proficiency, the Japanese speakers were

[30] The results discussed below were first reported in Brown (1996).

divided into two experimental groups: low-level (n=20) and high-level (n=15). The relevant background data are given in Table 5.

Table 5
Subject Information

Group	Mean Age at Testing	Mean Age of Exposure	Mean Years Studied
Low-level	19	11.7	7.6
High-level	24.5	12	11.5
Controls	35	—	—

Contrasts investigated & hypotheses

Two experimental contrasts were tested in this experiment, /l-r/ and /b-v/; the native /p-b/ contrast was also included as a control item. If perception and acquisition of non-native contrasts is constrained by the features of the native grammar, then since both beginning Japanese learners of English and more advanced Japanese learners of English have the same native grammar, they should both be able to perceive the /b-v/ contrast, yet unable to perceive the /l-r/ contrast.

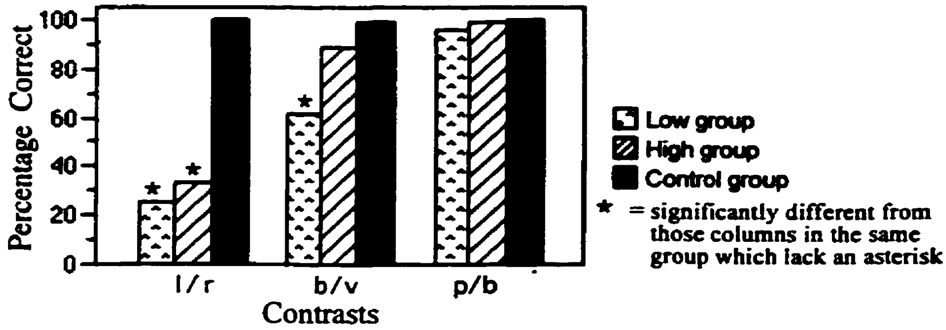
Tasks & materials

The tasks and materials used in this experiment were the same as those used in experiment 1; an AX Discrimination task was used to assess perception and a Forced Choice Picture Selection task was used to assess phonological competence.

Results & discussion

On the auditory discrimination task, a response that the two words in the minimal pair were "different" was counted as correct and a "same" response was counted as an error. Performance scores on each of the contrasts were tabulated separately for statistical analysis. Figure 7 reports the mean performance scores on each of the contrasts for each group.

Figure 7
Overall Auditory Performance by Contrast



From both the Low-level and High-level groups' near perfect performance on the control items (i.e. native /p–b/ pairs), it is clear that the task itself does not pose any difficulty for the learners. Furthermore, the control group performed as expected, accurately discriminating each of the three contrasts, with no significant difference between contrasts [F (9, 20) = 1.09, p = .36].

As the graph illustrates, both groups of Japanese speakers were significantly worse than the English controls at discriminating the /l-r/ contrast [F (2, 44) = 74.49 , p = .0001]. However, there was no difference between the Low-level and High-level groups in their ability to discriminate this contrast; learners in both groups were unable to perceive the difference between /l/ and /r/. Thus, an increase in English proficiency does not appear to affect perception of this non-native contrast. Accurate perception is blocked by the native grammar in the earliest stages of acquisition and continues to prevent perception even as the learner progresses.

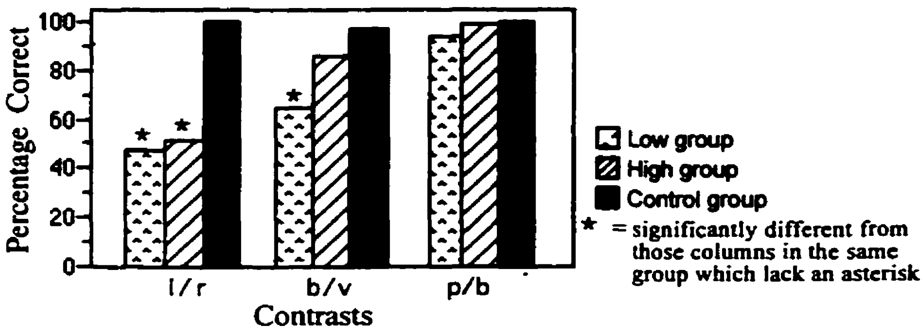
The situation is slightly different with respect to the other non-native contrast. While the learners in the Low-level group were not as accurate at discriminating the /b-v/ contrast as the learners in the High-level group, there was no difference between the High-level and the control groups' performance on this contrast [F (2, 44) = 9.79, p = .0003]. Thus, there was improvement in the Japanese speaker's ability to perceive this non-native contrast. We must keep in mind, though, that the Low-level group's somewhat poorer ability to discriminate /b/ and /v/ is still much better than either Japanese group's ability to distinguish /l/ from /r/. Finally, there was no statistical difference between the three groups in their ability to perceive the native /p-b/ contrast: all Japanese speakers discriminated this contrast as well as the native controls did [F (2, 44) = 1.08 , p = .35]. In short, whereas the ability to accurately perceive the /l-r/ contrast does not improve over time, the ability to perceive the /b-v/ contrast does improve, from being fairly good to being native-like.

We can now evaluate the relative effect of the native grammar on each of the contrasts at different stages of acquisition by examining the performance of each group individually. Beginning with the Low-level group, we find that their performance on each of the contrasts is significantly different from each other [F (19, 40) = 73.53 , p = .001]. That is, performance on the /p-b/ contrast, which is native-like, is better than performance on the /b-v/ contrast, which is better than performance on the /l-r/ contrast. In contrast, there was no difference in the High-level learners' ability to discriminate the /b-v/ and /p-b/ contrasts; both were perceived equally well and more accurately than the /l-r/ contrast [F (14, 30) = 91.75, p = .001]. These data show that at both stages of acquisition, the Japanese speakers are unable to discriminate /l/ and /r/, perceiving them, instead, as members of a single category. However, they differ in their ability to distinguish /b/ from /v/, indicating that the influence of the native grammar is not static, but changes as the learner's interlanguage grammar develops.

To summarize, these data allow us to see the influence of the native grammar at different stages of acquisition. We found that the ability to discriminate the /l-r/ contrast does not change over time, whereas learners do improve in their ability to perceive the /b-v/ contrast. We might be tempted to conclude from this that the influence of the native grammar simply changes over time, constraining perception more tightly in the early stages of acquisition but gradually weakening as the learner's interlanguage grammar develops. However, the situation is a bit more complex. We know that the native grammar influences perception of non-native sounds in two ways: it may either *block* perception or *facilitate* perception, depending on whether the relevant feature is present or absent in the L1 grammatical system. Looking at the data again, we see that perception of the /l-r/ contrast does not improve; it is only the perception of /b-v/ which improves. Thus, when the relevant feature is absent from the native grammar, as it is in the case of /l-r/, and perception is blocked, the effect of the grammar remains constant. However, if the relevant feature is present in the native grammar, as it is for /b-v/, then the effect of the grammar may change. In other words, the negative influence of the native grammar on perception is absolute, but the positive influence of the native grammar is enhanced as the learner progresses.

For the picture identification task, selection of the target picture was counted as a correct response and selection of the contrast picture was counted as an error. The groups' overall performance is compared in Figure 8.

Figure 8
Overall Picture Performance by Contrast



Near perfect performance was attained by the Japanese group on the control items in this task. Again, the performance of the control subjects – accurate and with no differences between contrasts – ensures that our task and materials are reliable [F (9, 20) = 2.39, p = .12].

The pattern of performance on this picture task is very similar to that on the auditory task. Both groups of Japanese speakers were significantly worse than the English controls at distinguishing lexical items that differed by /l/ or /r/ [F (2, 44) = 35.20, p = .0001]. Yet, there was no difference between the Low-level and High-level groups in their ability (or inability) to distinguish this contrast; learners in both groups were unable to discriminate /l/ and /r/ phonologically. This indicates that neither the beginner learners nor the more advanced learners have acquired the phonological structure necessary to differentiate these segments in their interlanguage grammars. This is not the case, though, with the /b-v/ contrast. While the learners in the Low-level group were not as accurate as the learners in the High-level group at distinguishing items that differed by /b/ or /v/, there was no difference between the High-level and the control groups' performance on this contrast [F (2, 44) = 5.43, p = .007]. Thus, duplicating the results from the auditory task, the ability to differentiate /b/ and /v/ in one's interlanguage grammar appears to develop over time. This suggests that there are, in fact, stages of phoneme acquisition. With respect to the native /p-b/ contrast, there was no statistical difference between the three groups: all Japanese speakers discriminated this contrast as well as the native controls did [F (2, 44) = 1.94 , p = .15].

Looking at each group individually, we find that the performance of the Low-level group on each of the contrasts is significantly different from the others [F (19, 40) = 47.81 , p = .0001]. We find the same pattern of performance by the High-level group [F (14, 30) = 31.82, p = .0001]. Both groups are better at distinguishing the /b-v/ contrast than they are the /l-r/ contrast, but they are still not as good at distinguishing the /b-v/ contrast as they are their own native /p-b/ contrast. It is clear from the data that the learners do not differentiate /l/ and /r/ in lexical items (i.e. the same structure is used to represent both segments). It is also clear that they do have distinct representations for /p/ and /b/.

What, then, is the status of the /b-v/ contrast, which seems to fall somewhere between the other contrasts; has it been acquired or not? I think the answer to this question is different for the two groups. In the case of the Low-level group, it appears that the new representations have not been acquired. This is not so surprising, given their perception of the /b-v/ contrast, which while quite good is not native-like. It is possible that these learners have not yet detected contrastive use of these segments in English and, as a result, have not yet acquired the new representations. In the case of the High-level group, however, I think we can be confident that they have acquired the new representations. Importantly, their perception of /b-v/ is native-like; thus a necessary condition for proper acquisition has been met. Moreover, their ability to distinguish /b/ and /v/ in this task (although poorer than their ability to distinguish the /p-b/ contrast) is as good as the native speakers' ability to distinguish /b/ and /v/, who undoubtedly differentiate these two sounds in their grammars.[31]

The research question we attempted to answer in this experiment was whether the influence of the native grammar on the perception and acquisition of non-native contrasts changes over time as the L2 learner progresses. The data demonstrate that there is not one answer to this question. The effects of the grammar, either to *block* or to *facilitate* perception and acquisition are differentially altered by the learner's development. If the feature underlying a non-native phonemic contrast is absent from the native grammar, then the native phonological system will continue to funnel the distinct acoustic signals for those sounds into a single perceptual category throughout the learner's development; perception of these non-native contrasts will not improve. In this case, the influence of the native grammar is rigid, immutable by increased exposure to a second language. If, however, the feature underlying a non-native phonemic contrast is present in the native grammar, the capacity of the native phonological system to use this feature in the processing of non-native sounds will be enhanced over the course of the learner's development; perception of these non-native contrasts will improve. In this case, increased exposure to a second language will actually strengthen the facilitative influence of the native grammar.

4.4 Summary & Conclusions

The goal of the research program presented in this chapter is to develop a comprehensive theory-driven model of L2 phoneme acquisition which accounts for the interrelation between perception and phonological acquisition and explains how and why this interrelation affects L2 phonological acquisition. It was proposed that the

---

[31] Depressed performance by both the High-level and the native control group on the /b-v/ contrast is likely caused by acoustic properties of this pair of sounds, especially in the environment of high front vowels, which minimizes their distinctiveness.

monotonic acquisition of feature geometric structure by young children reduces their perceptual sensitivity to particular non-native contrasts and that this adult feature geometry continues to mediate between the acoustic signal and the linguistic processor in adult speech perception, constraining which non-native contrasts adult learners will be sensitive to in the L2 input and, therefore, capable of acquiring. Thus, in order to fully understand why the L1 grammar exerts such a profound influence on the perception and acquisition of non-native phonemic contrasts, it is important that we understand the development of these systems in first language acquisition and their operation in the mature speaker.

Having determined how the interrelation between speech perception and the phonological system originates, we are in a better position to capture the nature of the mechanism that maps the L2 input onto L1 phonological categories; utilizing the tools of Feature Geometry theory enables us to formally articulate this mapping process. The central claim of the theory of phonological interference developed here is that the L1 influence found in L2 phonological acquisition is a consequence of how the speech perception mechanism operates in the native speaker. Based on the proposal that the decline in infants' ability to discriminate non-native contrasts is caused by the acquisition of phonological structure, speech perception in the native speaker will continue to be constrained by phonological properties of his or her native language throughout adulthood; more specifically, all speech sounds (native and non-native) will be perceived in terms of the features exploited by that particular language.

The experimental studies reported above have demonstrated that not all non-native contrasts are created equal: learners with the same L1 have more difficulty perceiving and acquiring some non-native contrasts than they do others; likewise, certain non-native contrasts are easily perceived and acquired by speakers of some languages, while those same contrasts will not be perceived or acquired by speakers of other languages. These differences, both between contrasts and between speakers of different languages, were argued to follow directly from the status of the relevant distinctive feature in the learner's L1 grammar: presence of the contrastive features in the grammar serves to sort the acoustic signal along that particular dimension, mapping the signals for two segments onto distinct phonological categories, whereas absence of the contrastive feature entails that the acoustic signals for the phonemes be mapped onto a single phonological category.

We saw in experiment 1 that Japanese learners' perception and acquisition of various English contrasts differed in exactly this respect: they were able to perceive and, therefore, acquire those contrasts which are distinguished by a feature that their native grammar employs for independent reasons (e.g., /b–v/ and /f–v/), but were unable to perceive that contrast that is distinguished by a feature not utilized in the L1 (e.g., /l–r/). Similarly, experiment 2 provided evidence that Japanese, Korean and Chinese speakers differ in their acquisition of particular English contrasts just as the model predicts: speakers of all three languages were unable to perceive the /s–θ/ contrast, which relies on a feature absent from all three L1s, yet were able to perceive those contrasts distinguished by features present in all three L1s (i.e., /p–f/ and /f–v/); most importantly, speakers of these three languages differed in their ability to perceive and acquire precisely that contrast, /l–r/, which is distinguished by a feature whose status differs among these languages.

One implication of my proposal is that prior to the development of a phonological system, infants should be able to perceive contrasts which they will fail to perceive as adults. With respect to Japanese speakers, if their difficulty discriminating /l/ and /r/ does indeed stem from the interference of their phonological system (rather than to, say, some genetic property of Japanese speakers), then before that system is in place, accurate perception of those sounds should be possible. Japanese infants have, in fact, been shown to perceptually distinguish /l/ from /r/ (Tsushima et al., 1994). Thus, the inability of Japanese speakers to discriminate /l/ and /r/ as adults does indeed appear to be a consequence of language development.[32] Moreover, an early study by

[32] Cochrane (1980) demonstrates that preadolescent Japanese children (ages 3–13 years) were no better than adults at perceiving /l–r/ minimal pairs. This finding indicates that the inability of the Japanese adults to perceive this contrast is a result of a change that occurs very early in language development (i.e., acquisition of phonological structure) and not the result of a more general change that occurs sometime prior to puberty (e.g., lateralization of brain function).

Miyawaki et al. (1975) shows that the difficulties Japanese speakers have discriminating /l/ and /r/ are due to specific properties of their perception of *speech*, not to deficiencies in their basic auditory mechanisms. These researchers found that adult Japanese speakers accurately discriminate /l/ and /r/ when they are presented in a "non-speech mode".[33] In other words, Japanese speakers are able to discriminate /l/ and /r/ when the acoustic signal is processed directly by the auditory system, rather than the linguistic module.

We know from first language acquisition research that the development of segmental structure involves the interaction of Universal Grammar and the learner's detection of phonemic contrasts in the input. Thus, successful acquisition of novel phonemes by L2 learners depends not only on the availability of UG, but, importantly, on adequate intake to the language acquisition device. By demonstrating that some L2 learners do not perceive the L2 input correctly (in fact, precisely those learners who are unable to acquire the given contrast), this research strongly suggests that the inability of some L2 learners to acquire novel phonemic contrasts is due to the lack of proper input, rather than the unavailability of UG. Thus, the failure of L2 learners to acquire novel phonemes should not necessarily be taken as evidence that UG is not available in L2 acquisition. In fact, these results demonstrate that if L2 learners are able to perceive a non-native contrast, they are able to acquire that contrast, suggesting that the mechanism for constructing novel segmental representations (which is arguably part of UG) is still operative in L2 acquisition.

These findings fit in nicely with recent trends in second language acquisition theory which suggest that differences in L1 and L2 acquisition (as well as differences across learners in L2 acquisition) stem not from the inavailability of Universal Grammar but rather from the initial state of acquisition (papers in Schwartz & Eubank, 1996). The goal of this new line of research is to define the initial state of L2 acquisition and, thereby, explain the development of the L2 grammar. Differences in L2 acquisition of a

---

[33] In the "non-speech mode", all of the acoustic information that does NOT differentiate the two sounds - namely the first and second formants - was removed from the stimuli, resulting in something that sounds like a high-pitched glissando.

particular language that covary with learners' native language are now assumed to be a result of the L2 initial state. The research agenda, then, is to define this initial state. Research on the acquisition of syntactic properties of the L2 has been used to support several hypotheses regarding the linguistic content of the initial state (see Schwartz & Eubank, 1996, for Schwartz & Sprouse's *Full Transfer/Full Access Model*, Vainikka & Young-Scholten's, *Minimal Tree Hypothesis*, and Eubank's *Weak Transfer Hypothesis*). The findings from experimental research on the L2 acquisition of phonemes seem to be most consistent with Schwartz & Sprouse's hypothesis that the entire L1 system forms the initial basis of L2 acquisition: all of the data indicate that in the earliest stages of L2 acquisition, L2 phonemes are mapped according to the L1 feature geometry onto L1 phonemic categories; only subsequently are new L2 categories acquired.

The claim that the entire L1 phonological system constitutes the initial state for L2 phoneme acquisition raises an interesting question: If the acoustic signal is perceived in terms of the learner's L1 phonemic categories, how can the learner accurately perceive non-native sounds in order for new phonemic categories to be established? In other words, how can the input be mapped by the adult feature geometry onto new L2 categories when those categories don't yet exist? The answer to this question, I believe, depends on whether it is the *phonemes* or the *features* of the L1 which constrain perception. If the acoustic signal is mapped onto L1 phonemes, then it would seem that it is the phonemic level which impinges upon L2 acquisition. Yet, I have argued that it is the featural level which is relevant.

A closer examination of the data from low and high proficiency learners presented in experiment 3 suggests how these two positions might be reconciled. In particular, these data suggest that initially, in the earliest stages of acquisition, the phonemes of the L1 have a profound influence on the perception of non-native contrasts. In an attempt to understand the L2 input, and in the absence of new phonological categories, the L2 input is fitted into the L1 system any way it can be (often by brute force, ignoring variations that the system senses but cannot yet deal with appropriately). For example, the acoustic signals for both /b/ and /v/ will be mapped by Japanese speakers onto the L1 phoneme /b/ and the acoustic signals for /l/ and /r/ will be mapped onto /ɾ/. This will be the initial stage of acquisition and, incidentally, the situation for loanword phonology.[34]

However, despite the initial attempt of the L1 system to accommodate all of the input within L1 structures, portions of the L2 input will not map adequately to the L1 system, as a result of the presence of the relevant contrasting feature. Taking our example again, the English /b/ will map completely onto the Japanese phoneme /b/, but the English /v/ will not map precisely to the Japanese category.[35] Thus, although both are perceived as a single category in the early stages of acquisition, the presence of the feature [continuant] ensures that the acoustic signal for /v/ does not correspond exactly to a Japanese category; this slight mismatch between the L2 input and the L1 structures will cause perceptual reorganization (the beginner learners in experiment 3). Over the course of development, and with increased exposure, a new phonological category will be established; following the establishment of this new category, the original native category will be by-passed entirely in perception, and perception of those contrasts will be native-like (the advanced learners in experiment 3). If, however, the feature that distinguishes a given non-native contrast is absent from the L1 grammar, then the L2 input will map perfectly to an existing L1 category and there will be no trigger for

---

[34] Loan words in Japanese are written in katakana, one of the two Japanese syllabary writing systems. When a foreign word containing the segment [v] is borrowed, this segment is traditionally transcribed as one of the kana for [ba], [bi], [be], [bo] or [bu] (i.e., [b] is substituted for [v]). However, within the last five years, a new kana symbol has been introduced by *Monbusho* (Japan's Ministry of Education) to represent the sound [v], ヴ, which is the symbol for the vowel [u], plus voicing marks). Thus, just as it is possible for the learner, who originally maps all L2 sounds onto L1 categories (even those that do not match perfectly), to acquire a new perceptual category for those L2 sounds that do not correspond perfectly to the L1 categories, so to can writing systems be adapted to better represent the original pronunciations of loanwords. Words that were borrowed into Japanese before the introduction of this new symbol for [v] continue to be written as though they contained a [b] (e.g., "boɾiboɾu" for volleyball), but words that have been borrowed after the introduction of this symbol are written to accurately reflect the language of origin's pronunciation (e.g. "Bon Jovi" for Bon Jovi). It is quite interesting (and not accidental, I think) that a new katakana symbol has been introduced for [v] (an L2 phoneme which Japanese speakers have been shown above to accurately perceive), but not for [l] (an L2 phoneme which these speakers do not accurately perceive). That the writing system has been modified to accommodate [v], but not [l], reflects, I think, the increasing perceptual awareness of Japanese speakers that [v] does not adequately correspond to any native Japanese phonemes.

[35] This position predicts that the time required to process English /b/ and /v/ by Japanese speakers, for example, will differ. Since English /b/ maps exactly to the Japanese category, it should be identified as /b/ more quickly than /v/ is identified (as /b/ or /v/).

acquisition, as was the case with both our beginner and advanced learners for the /l–r/ contrast. Thus, while the input is initially sorted in terms of L1 phonemes, it is the L1 features which guide this mapping process and, therefore, determine to what extent the L2 input can be accommodated by existing phonological structure; in this way the features also constrain which non-native contrasts will be acquired by the learner.

In 1939, Trubetzkoy wrote that "once a speaker has learned to attend to certain features only, he supposedly approaches all other languages through his own 'grid' of distinctive versus non-distinctive features" (p. 54). Although our theories of segmental representation and language acquisition have evolved since then, the three experimental studies discussed in this paper demonstrate that Trubetzkoy was essentially correct: a speaker's L1 grammar constrains the non-native contrasts that he or she will be able to accurately perceive and subsequently acquire. The theory of phonological interference developed in this paper helps us to better understand *why* and *how* the L1 exerts an influence on the acquisition of a second phonological system. By isolating those aspects of the L1 grammar which are responsible for this influence (i.e., the phonological features) as well as providing an explanation of the mechanisms which exert this influence (i.e., the operation of speech perception in the monolingual), we can not only explain differences in a learner's ability to acquire various non-native contrasts, we can also account for the varying abilities of learners with different L1's to acquire the same non-native contrast.

Conclusions to the Thesis

The research that has been presented in this thesis is very much theoretically driven: assuming speakers' knowledge of phonemes to consist of feature-geometric representations in their mental grammars has influenced how we understand the acquisition of that knowledge by children, and how we account for the use of that knowledge by mature speakers in their perception of speech sounds. While our theories of acquisition are dependent upon our theories of grammar, satisfactory explanation of the acquisition and use of linguistic knowledge in terms of the constructs of a particular theory also lends (indirect) support for those very constructs.

To conclude this dissertation, I will review the major results of the research presented here and consider how these findings contribute to the goals of linguistic theory, as well as to phonological theory and our theories of acquisition. Three of the primary goals of a generative theory of grammar are to: 1) formally characterize the knowledge that speakers have about their native language, 2) determine how that knowledge (i.e., the constructs posited to represent that knowledge mentally) is acquired, and 3) ascertain how that knowledge (again, mental constructs) is used in the production and comprehension of language (Chomsky, 1981, 1986). As each of the four papers in this thesis has been previously summarized, and conclusions from each have been drawn in the separate chapters, I will not spend unnecessary time and paper (re)summarizing the research in each one at this time; rather, I will focus on how the general findings of all four combine to provide some answers to these three broad issues.

Let us first consider the goal of determining what constitutes knowledge of phonemes. I have assumed in this thesis that speakers' knowledge of phonemes (i.e., their behavior in phonological processes and their contrastive nature within an inventory) is best characterized by feature-geometric representations, as put forth in Feature Geometry theory. According to this theory, the relations amongst features are captured structurally in terms of hierarchical arrays of those features within the segment, at all levels of the phonology (see Appendix E for a comparison of this theory with Optimality Theory). The main results of each component of research presented in this thesis support this position.

Chapter 1 presented evidence that even the mysterious behavior of both sonorant and non-sonorant lateral segments could be captured by feature geometric representations. Particularly at issue was whether the feature [lateral], traditionally assumed to distinguish lateral segments from non-lateral segments, should be included in the universal hierarchy of phonological features. The arguments discussed there suggest that [lateral] is not a phonological feature and that laterality (in both sonorant and non-sonorant laterals) is a phonetic property derived from specific feature-geometric representations. Until now, laterals and their representation have resisted straightforward treatment in Feature Geometry theory. By demonstrating that the full range of phonological behavior of lateral segments can be captured in terms of feature-geometric representations, I deflect the criticism that Feature Geometry is inadequate as a theory of segmental representation. Thus, this research provides support for the claim that a speaker's knowledge of phonemes is best characterized in terms of the hierarchical organization of features within the segment; it also pinpoints what constitutes knowledge of laterality, in particular that [lateral] is not part of a speaker's phonological knowledge of segments and that [Coronal] differentiates the lateral and central approximants /l/ and /r/.

We turn now to consider the second goal of linguistic theory: to determine how this knowledge of phonemes is acquired. Of course, when we investigate the acquisition of knowledge of phonemes (or any other type of linguistic knowledge, for that matter), what we are actually doing is trying to determine whether (and how) the mental constructs that linguists have posited to characterize that knowledge can be acquired. Thus, to the degree that acquisition research can experimentally demonstrate that feature-geometric representations can be acquired, as constrained by the theory of Feature Geometry posited in UG, it provides (indirect) support for the claim that feature-geometric representations are the correct characterization of that knowledge.

The main result of Chapter 2 was that knowledge of phonemes of the ambient language is gradually acquired; that is, segmental representations are progressively elaborated, in a systematic order, until all of the target language's phonemes are differentiated in the emerging grammar. This systematic order of acquisition, as well as the particular order observed, was argued to stem from the active involvement of Feature Geometry in phoneme acquisition: as a theoretical characterization of innate properties of the developing human nervous system (i.e., part of UG), Feature Geometry provides the featural material and constrains the process by which children acquire phonemic representations. By establishing how children acquire segmental representations (and showing that there is a way into the developmental state and a way out), this research demonstrates that feature-geometric representations are indeed acquired on the basis of positive evidence and that knowledge of the hierarchical arrangement of features is a component of UG, thereby providing indirect support for the theory of Feature Geometry itself.

Finally, with respect to the third goal, of ascertaining how linguistic knowledge underlies language use, Chapters 3 and 4 outlined the way in which phonological knowledge impinges upon the comprehension and processing of speech sounds. Here again, the demonstration that the feature-geometric representation of phonemes can aid our understanding of the influence that phonemic knowledge exerts on the speech perception system also strengthens the claim that this knowledge is mentally represented in terms of hierarchical feature arrays. Our examination of the development of speech perception and L1 phoneme acquisition indicated that a child's perceptual capacities worsen as his or her ability to discriminate phonemes structurally in the grammar improves. This was taken as evidence that the elaboration of Feature Geometry in the child's grammar imposes the boundaries according to which phonemes will be categorized perceptually.

Thus, not only does the acquired hierarchy of features reflect which segments are contrastive in a speaker's grammar, it also corresponds to properties of the speech perception system. Thus, the research presented in Chapter 3 and Chapter 4 on the perception and acquisition of several non-native phonemic contrasts by Japanese, Korean and Chinese speakers suggests that feature-geometric representations (once acquired) have psychological consequences in the perception of speech sounds, lending strong support to the view that a speaker's knowledge of phonemes is hierarchical in nature and that segmental representations correspond to psychological properties of the human mind.

In the remainder of the conclusion, I turn to consider what these findings might contribute to phonological theory, second language acquisition theory and our conception of UG. Let us start with phonological theory. Although most phonological theories, including Feature Geometry theory, recognize that lexical items are decomposable into segments which are themselves decomposable into features, very few (if any) would accord independent status to all three. That is, while phonologists talk of a language's *inventory* of phonemes or *inventory* of features, these are not viewed as existing independently from the lexical items that they comprise; rather the "inventory" itself is epiphenomenal (i.e., it is the aggregate of features or segments present in lexical items). The acquisition and perception data reported in this thesis, however, suggest that there may be evidence for features and segments as autonomous types of phonological knowledge.

In the theory of phoneme acquisition outlined in Chapter 2, it was tacitly assumed (though not explicitly stated) that the feature hierarchy exists separately from the individual lexical items that incorporate those hierarchically organized arrays. In particular, it was argued that UG provides the emerging grammar with a minimal amount of Feature Geometry which is subsequently elaborated on the basis of detection of contrasts in the input. Now, this Feature Geometry in the emerging grammar could either be interpreted as simply a theoretical model of the amalgamation of individual phonemes' feature geometries or it could be viewed as an independent linguistic entity which represents a different level (or type) of phonological knowledge.

The findings from the perception of non-native sounds suggests that this Feature Geometry does, in fact, have an independent status and is separate from representations of individual segments or lexical items. The evidence for this is L2 learners' ability to perceive non-native contrasts, as facilitated by the features utilized in their grammar. Learners are not restricted to perceiving only the feature patterns instantiated in their native segmental representations; rather, the adult feature geometry would appear to operate independently of native phonemes or lexical items. Thus, it is not the L1 representations themselves, but the components of those representations which map the acoustic signal onto native perceptual/phonemic categories.

We find similar evidence for the independent status of segments in the grammar. It has been difficult to find physical or psycholinguistic evidence for the segment; the fact that it is very difficult to match the overlapping acoustic signal to discrete segments and that phoneme awareness seems to be an artifact of reading education has been a factor in dissuading phonologists from positing a phoneme inventory in the grammar that is separate from the lexical items of which the phonemes are a part. However, the perceptual findings from the studies reported here suggest that there is indeed an independent phoneme inventory in the grammar of speakers. Although it is the feature geometry which guides the mapping, the acoustic signal is mapped onto L1 phonemic (or perceptual) categories. Thus, L2 sounds are perceived and identified as instances of L1 segments.[1] This role of the L1 phonemes is most evident in the early stages of L2 phoneme acquisition in which all of the L2 input are mapped onto these L1 categories; only subsequently will new categories be established for certain L2 segments.

Moving now to acquisition theory, the findings regarding what triggers the acquisition of segmental representations is very important. We have established that it is

[1] Quite interestingly, this is the case for Japanese and Chinese speakers, both of whom do not have phonemic writing systems. Learners who have syllabic and ideographic scripts, nevertheless, show psycholinguistic evidence of having perceptual categories that correspond to phonemes.

the detection of a contrast in the input that causes the acquisition of phonemic contrasts. Quite broadly, this knowledge could be applied to different teaching situations (either in speech therapy context or language teaching context), in which the crucial task of the instructor would be to heighten awareness, either through different activities or manipulation of the input, of the contrastive nature of segments. But even more significantly, this research demonstrates that when acquisition researchers investigate L2 acquisition (of phonological knowledge or any other type of linguistic knowledge) it is imperative to establish that the learners are receiving the appropriate input. "Input" here is not meant in the traditional sense of the relevant input being present in the language environment, but rather as intake to the acquisition device. If we are to evaluate the acquisition of L2 knowledge, then we need to verify, experimentally if possible, that the learner is perceiving (or analyzing) the L2 input in a way that is faithful to the phonetic, syntactic, morphological or semantic information contained in that input. Successful acquisition of L2 structures depends not only on the availability of UG, but, crucially, on accurate perception and analysis of the input. We may ultimately determine that learners cannot acquire properties of the L2 that are governed by UG precisely because they are not able to extract the relevant information from the input, due to the interference of their L1 grammar; but this, I believe, is different from concluding that UG is inaccessible.

Finally, this brings us to the nature of the role of the L1 grammar in second language acquisition. It has become increasingly clear, from the results of these studies and many others, that the L1 exerts a tremendous influence on the course of L2 acquisition. Many have concluded from these studies that UG is only accessible in L2 acquisition via one's L1 grammar. There two possible ways in which the L1 might constrain L2 acquisition that would create the impression that only properties of UG instantiated in the L1 may be acquired; each takes a different view on the relationship between UG and the L1 grammar. The first alternative is that UG is accessible via the L1 precisely because UG was, in fact, transmuted into the L1 grammar in the course of

L1 acquisition (see e.g., Chomsky & Lasnik, 1993). Under this view, acquisition of L2 structures or principles should be limited to those that are also found in the L1.

The other alternative is that access to UG appears to be limited by the L1 because the native grammar (in many instances) disrupts the flow of information to UG that would otherwise trigger the acquisition of L2 structures. Under this view, much of the relevant L2 input will be filtered out by the L1 grammar, but in those limited instances in which it was not, acquisition of L2 properties should be possible. All of the L2 studies discussed in this dissertation demonstrate that when learners perceive a novel contrast in the L2 input, new perceptual categories and phonological representations can be acquired. Thus, the data presented here argue in favor of the view of UG and L1 acquisition in which properties of UG are transmitted to the emerging grammar, but that UG itself remains intact, providing L2 learners information about the L2 not deducible from the input when the conditions necessary for acquisition are met.

References

Abramson, A. & Lisker, L. (1970) Discriminability along the voicing continuum: Cross-language tests. *Proceedings of the Sixth International Congress of Phonetic Sciences*, pp. 563-567. Prague: Academia.

Anderson, J. & Ewen, C. (1987) *Principles of Dependency Phonology*. Cambridge: Cambridge University Press.

Aoki, H. (1970) *Nez Perce Grammar*. Berkeley: University of California Press.

Archangeli, D. (1988) Aspects of Underspecification Theory. *Phonology* **5**(2):183-207.

Archangeli, D. (1984) *Underspecification in Yawelmani Phonology and Morphology*. Doctoral Dissertation. Massachusetts Institute of Technology. [Published 1988, Garland, New York].

Archibald, J. (1993) *Language Learnability and L2 Phonology: The Acquisition of Metrical Parameters*. Dordrecht: Kluwer Academic Publishers.

Aslin, R., Pisoni, D., Hennessy, B. & Perry, A. (1981) Discrimination of voice onset time by human infants: New findings and implications for the effect of early experience. *Child Development* **52**:1135-1145.

Avery, P. (1993) The specification of laryngeal features. Paper presented at the Quebec–Ontario Phonology Workshop. Montreal, Canada, February, 1993.

Avery, P. & Rice, K. (1989) Segment structure and coronal underspecification. *Phonology* **6.2**: 179-200.

Barker, M. (1964) *Klamath Grammar*. Berkeley: University of California Press.

Barton, D. (1975) Statistical significance in phonemic perception experiments. *Journal of Child Language* **2**:297-298.

Barton, D. (1976) *The Role of Perception in the Acquisition of Phonology*. Doctoral dissertation, University of London, England. Distributed by the Indiana University Linguistics Club.

Barton, D. (1980) Phonemic perception in children. In G. Yeni-Komshian, J. Kavanagh & C. Ferguson (eds.) *Child Phonology*, Vol 2, pp. 97-116. New York: Academic Press.

Beddor, P. & Gottfried, T. (1995) Methodological issues in cross-language speech perception research with adults. In W. Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, pp. 207–232. Baltimore: York Press.

Bellin, W. (1984) Welsh phonology in acquisition. In M. Ball & G. Jones (eds.) *Welsh PhonologY*, pp. 158-175. University of Wales Press: Cardiff.

Burnham, D. (1986) Developmental loss of speech perception: exposure to and experience with a first language. *Applied Psycholinguistics* **7**:207-240.

Berwick, R. (1985) *The Acquisition of Syntactic Knowledge*. Cambridge, Ma.: MIT Press.

Best, C. (1994) The emergence of native-language phonological influence in infants: A perceptual assimilation model. In H. Nusbaum, J. Goodman & C. Howard (eds.) *The Transition from Speech Sounds to Spoken Words: The Development of Speech Perception*, pp. 167-224. Cambridge, Mass.: MIT Press.

Best, C. (1993) Emergence of language–specific constraints in perception of non-native speech: A window on early phonological development. In B. de Boysson-Bardies (ed.) *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, pp. 289-304. Dordrecht: Kluwer Academic Publishers.

Best, C. & McRoberts, G. (1989) Phonological influences in infants' discrimination of two non-native speech contrasts. Paper presented at the Society for Research in Child Development, Kansas City, Kansas, April 1989.

Best, C., McRoberts, G. & Sithole, N. (1988) Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* **14**: 345-360.

Bley-Vroman, R. (1989) The logical problem of second language acquisition. In S. Gass & J. Schachter (eds.) *Linguistic Perspectives on Second Language Acquisition*, pp. 41-72. Cambridge: Cambridge University Press.

Bond, Z. & Adamescu, L. (1979) Identification of novel phonetic segments by children, adolescents and adults. *Phonetica* **36**:182-186.

Borden, G., Gerber, A. & Milsark, G. (1983) Production and perception of the /r/–/l/ contrast in Korean adults learning English. *Language Learning* **33**: 499-526.

Braine, M. (1971) On two types of models of the internalizations of grammars. In D. Slobin (ed.) *The Ontogenesis of Grammar*, pp. 153-188. New York: Academic Press.

Brière, E. (1966) An investigation of phonological interferences. *Language* **42**: 768-795.

Broselow, E. and Finer, D. (1991) Parameter setting in second language phonology and syntax. *Second Language Research* **7**:35-59.

Broselow, E. & Park, H.B. (1995) Mora conservation in second language prosody. In J. Archibald (ed.) *Phonological Acquisition and Phonological Theory*, pp. 151-168. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Browman, C. & Goldstein, L. (1986) Towards an articulatory phonology. *Phonology Yearbook* **3**:219-252.

Browman, C. & Goldstein, L. (1989) Articulatory gestures as phonological units. *Phonology* **6**:201-251.

Brown, C. (to appear a) The interrelation between speech perception and phonological acquisition from infant to adult. In J. Archibald (ed.) *Second Language Grammars*. Oxford: Blackwell Publishers.

Brown, C. (to appear b) The role of the L1 grammar in the acquisition of L2 segmental structure. *Second Language Research.*

Brown, C. (to appear c) Accounting for cross-language differences in the perception and acquisition of second language speech sounds. *New Sounds 97, Proceedings of the Third International Symposium on the Acquisition of Second-Language Speech.*

Brown, C. (1996) English tongues, Japanese ears: A theory of phonological interference. *Language & Culture: Journal of the Institute of Language & Culture Studies* **30**:165-203. Hokkaido University, Sapporo, Japan.

Brown, C. (1995) The feature geometry of lateral approximants and lateral fricatives. In van der Hulst, H. and van de Weijer, J. (eds.) *Leiden at Last*, pp. 41-88. University of Leiden, Holland.

Brown, C. (1994) When the native grammar facilitates acquisition of non-native phonological contrasts. Paper presented at the Second Language Research Forum, Montréal, Canada. October, 1994.

Brown, C. (1993a) The role of the L1 grammar in the L2 acquisition of segmental structure.*McGill Working Papers in Linguistics* **9**:180-210.

Brown, C. (1993b) On the representation of laterality.*McGill Working Papers in Linguistics* **8.2**:109-142.

Brown, C. & Matthews, J. (in press) The role of feature geometry in the development of phonemic contrasts. In S. J. Hannahs & M. Young-Scholten (eds.) *Generative Studies in the Acquisition of Phonology* Amsterdam: John Benjamins Publishing Company.

Brown, C & Matthews, J. (1996) Explaining native language influence: How the native grammar may hinder or facilitate the acquisition of non-native contrasts. Paper presented at The Pacific Second Language Forum, Wellington, New Zealand.

Brown, C. & Matthews, J. (1993) The acquisition of segmental structure. *McGill Working Papers in Linguistics* **9**:46-76.

Brown, R. & Hanlon, C. (1970) Derivational complexity and the order of acquisition in child speech. In J. Hayes (ed.) *Cognition and the Development of Language*, pp. 11-54. New York: Wiley.

Burnham, D., Earnshaw, L. & Clark, J. (1986) Categorical identification of phonemic and non-phonemic bilabial stops by infants, children and adults. *Speech and Language Research Centre Working Papers* **4**:53-73.

Catford, J. & Ladefoged, P. (1968) Practical phonetic exercises. *UCLA Working Papers in Phonetics* **11**: 1-22. Los Angeles: UCLA Press.

Chao, Y.R. (1968) *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Chomsky, N. (1986) *Knowledge of Language: Its Nature, Origins and Use*. New York: Praeger.

Chomsky, N. (1981) *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Cambridge, Ma.: MIT Press.

Chomsky, N. & Lasnik, H. (1993) Principles and parameters theory. In J. Jacobs, A. von Stechow, W. Sternfeld & T. Venneman (eds.) *Syntax: An International Handbook of Contemporary Research*, pp. 506-569. Berlin: Walter de Gruyter.

Chomsky, N. & Halle, M. (1968) *The Sound Pattern of English*. Cambridge, Ma.: MIT Press.

Clahsen, H. & Muysken, M. (1986) The availability of universal grammar to adult and child learners: A study of the acquisition of German word order. *Second Language Research* **2**:93-119.

Clements, G.N. (1988a) The role of the sonority cycle in core syllabification.*Working Papers of the Cornell Phonetics Laboratory* **2**:1-68.

Clements, G.N. (1988b) Toward a substantive theory of feature specifications. Paper presented at the 18th Northeast Linguistics Society Conference (NELS).

Clements, G.N. (1985) The geometry of phonological features. *Phonology Yearbook* **2**:223-250.

Clements, G.N. & Hume, E. (1994) The internal organization of speech sounds. In J. Goldsmith (ed.) *A Handbook in Phonological Theory*, pp. 245-306. Oxford: Blackwell Publishers.

Cochrane, R. (1980) The acquisition of /r/ and /l/ by Japanese children and adults learning English as a second language. *Journal of Multilingual and Multicultural Development* **1**:331-60.

Cole, J. & Kisseberth, C. (1994) An optimal domains theory of harmony. ms., University of Illinois at Urbana-Champaign.

Corder, S. (1967) The significance of learners' errors. *International Review of Applied Linguistics* **5**: 161-170.

Curtin, S., Goad, H. & Pater, J. (to appear) On the acquisition of Thai VOT contrasts by native speakers of English. *Proceedings of the 3rd Generative Approaches to Language Acquisition Conference*. Edinburgh, Scotland.

Dinnsen, D. (1993) Underspecification and phonological disorders. In M. Eid & G. Iverson (eds.) *Principles and Prediction: The Analysis of Natural Language*, pp. 287-304. Amsterdam: John Benjamins Publishing Company.

Edwards, M. (1974) Perception and production in child phonology: The testing of four hypotheses. *Journal of Child Language* **2**:205-219.

Eilers, R., Gavin, W. & Oller, D. (1982) Cross-linguistic perception in infancy: Early effects of linguistic experience. *Journal of Child Language* **9**: 289-302.

Eilers, R., Oller, D. & Gavin, W. (1978) A cross-linguistic study of infant speech perception. Paper presented at the Southeastern Conference on Human Development: Atlanta, Georgia.

Eimas, P. (1975) Auditory and phonetic coding of the cues for speech: Discrimination of the /l–r/ distinction by young infants. *Perception & Psychophysics* **18**: 341-347.

Eimas, P., Siqueland, E., Jusczyk, P. & Vigorito, J. (1971) Speech perception in infants. *Science* **171**: 303-306.

Fallon, P. (1993) Liquid dissimilation in Georgian. *Proceedings of the Eastern States Conference on Linguistics (ESCOL)* **10**:105-116.

Fee, J. (1992) Vowel acquisition in Hungarian: Evidence for an order of feature acquisition. Paper presented at the 17th annual Boston University Conference on Language Development: Boston.

Ferguson, C. (1978) Learning to pronounce: the earliest stages of phonological development in the child. In F. Minifie & L. Lloyd (eds.) *Communication and Cognitive Abilities: Early Behavioral Assessment*, pp. 273-297. Cambridge: Cambridge University Press.

Ferguson, C. & Farwell, C. (1975) Words and sounds in early language acquisition. *Language* **51**:419-439.

Flege, J. (1995) Second language speech learning: Theory, findings, and problems. In W. Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, pp. 233–273. Baltimore: York Press.

Flege, J., Takagi, N. & Mann,V. (1995) Japanese adults can learn to produce English /ɹ/ and /l/ accurately. *Language and Speech* **38**: 25-55.

Flege, J. (1992) Speech learning in a second language. In C.A. Ferguson, L. Menn & C. Stoel-Gammon (eds.) *Phonological Development: Models, Research, Implications*. pp. 565-604. Baltimore: York Press.

Flege, J. (1991) Perception and production: The relevance of phonetic input to L2 phonological learning. In T. Heubner and C. Ferguson (eds.) *Crosscurrents in Second Language Acquisition and Linguistic Theory*, pp. 249-290. Philadelphia: John Benjamins Publishing Company.

Flege, J. (1981) The phonological basis of foreign accent. *TESOL Quarterly* **15**:443-55

Flynn, S. and Manuel, S. (1991) Age-dependent effects in language acquisition: An evaluation of "Critical Period" hypotheses. In L. Eubank (ed.) *Point/Counterpoint: Universal Grammar in Second Language Acquisition*. Amsterdam: J. Benjamins Publishing Company.

Fry, D., Abramson, A., Eimas, P., & Liberman, A. (1962) The identification and discrimination of synthetic vowels. *Language and Speech* **5**:171–189.

Garnica, O. (1973) The development of phonemic speech perception. In T. Moore (ed.) *Cognitive Development and the Acquisition of Language*, pp. 215-222. New York: Academic Press.

Garnica, O. (1971) The development of the perception of phonemic differences in initial consonants by English-speaking children: A pilot study. *Papers and Reports on Child Language Development* **3**:1-29. Stanford University Linguistics.

Gnanadesikan, A. (1995) Markedness and faithfulness constraints in child phonology. Ms., Rutgers University, New Brunswick, N.J. (ROA–67).

Goad, H. (1996) Consonant harmony in child language: Evidence against coronal underspecification. In B. Bernhardt, J. Gilbert & D. Ingram (eds.) *Preceedings of the UBC International Conference on Phonological Acquisition*, pp. 187-200. Somerville, Ma.: Cascadilla Press.

Goldsmith, J. (1993) Harmonic phonology. In J. Goldsmith (ed.) *The Last Phonological Rule*, pp. 21-60. Chicago: University of Chicago Press.

Goto, H. (1971) Auditory perception by normal Japanese adults of the sounds 'L' and 'R'. *Neuropsychologia* **9**: 317-323.

Guilfoyle, E. & Noonan, M. (1992) Functional categories and language acquisition. *Canadian Journal of Linguistics* **37**(2):241-272.

Haegeman, L. (1991) *Introduction to Government and Binding Theory*. Oxford: Blackwell Publishers.

Halle, M. (1983) On distinctive features and their articulatory implementation. *Natural Language and Linguistic Theory* **1**:91-105.

Hamilton, P. (1993) On the internal structure of the coronal node. *Proceedings of the Eastern States Conference on Linguistics (ESCOL)* **10**:129-140.

Hancin-Bhatt, B. (1994a) *Phonological Transfer in Second Language Perception and Production*. Unpublished Ph.D. Dissertation: University of Illinois at Urbana, Champaign.

Hancin-Bhatt, B. (1994b) Segment transfer: A consequence of a dynamic system. *Second Language Research* **10**(3):241-269.

Hockett, C. (1958) *A Course in Modern Linguistics*. New York: Macmillan.

Hualde, J. (1988) Affricates are not contour segments. *Proceedings of WCCFL* **VI**:77-89.

Hulst, H. van der (1989) Atoms of segmental structure: Components, gestures and dependency. *Phonology* **6**:253-284.

Hyams, N. (1987) The theory of paramters and syntactic development. In T. Roeper & E. Williams (eds.) *Parameter Setting*, pp. 1-22. Dordrecht: Reidel.

Hyams, N. & Wexler, K. (1993) Grammatical basis of null subjects in child language. *Linguistic Inquiry* **24**:421-459.

Ingram, D. (1989) Underspecification theory and phonological acquisition. Unpublished ms., University of British Columbia: Vancouver.

Ingram, D. (1979) Phonological patterns in the speech of young children. In P. Fletcher & M. Garman (eds.) *Language Acquisition*, pp. 133-149. Cambridge: Cambridge University Press.

Ingram, D. (1978) Sensori-motor intelligence and language development. In A. Lock (ed.) *Action, Gesture and Symbol: The Emergence of Language*, pp. 261-290. London: Academic Press.

International Phonetic Association (Principles of the) (1979) University College, London: International Phonetic Association.

Itô, J., Mester, A. & Padgett, J. (1995) Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* **26**(41):571-614.

Jakobson, R. (1941). *Kindersprache, Aphasie und Allgemeine Lautgesätze*. Uppsala: Almqvist & Wiksell. Citations from the English translation, *Child Language, Aphasia and Phonological Universals*, translated by A. Keiler, 1968. The Hague: Mouton.

Jakobson, R., Fant, G. & Halle, M. (1952) *Preliminaries to Speech Analysis*. Cambridge: MIT Press.

Jung, M-W. (1962) *A Contrastive Study of English and Korean Segmental Phonemes with Some Suggestions Toward Pedagogical Application*. M.S. Dissertation, Georgetown, University.

Kari, J.(1976) *Navho Verb Prefix Phonology*. New York: Garland Publishing, Inc.

Kawakami, S. (1977) *An Outline of Japanese Sounds [Nihongo Onsei Gaisetsu]*. Tokyo: Ouhussha.

Kawasaki, T. (1997) *Coda Constraints*. Ph.D. Dissertation: McGill University, Montréal, Canada

Kean, M.L. (1975) *The Theory of Markedness in Generative Grammar*. Ph.D. dissertation, MIT.

Kellerman, E. (1977) Toward a characterization of the strategy of transfer in second language learning. *Interlanguage Studies Bulletin* **2**(1): 58-145.

Kenstowicz, M. (1994) *Phonology in Generative Grammar*. Oxford: Blackwell Publishers.

Kiparsky, P. & Menn, L. (1977) On the acquisition of phonology. In J. McNamara (ed.) *Language, Learning and Thought*, pp. 47-78. New York: Academic Press.

Kuhl, P. (1979) Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America* **66**:1669-1679.

Kuhl, P. & Miller, J. (1975) Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science* **190**:69-72.

Kuhl, P. & Padden, D. (1983) Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *Journal of the Acoustical Society of America* **73**:1003-1010.

Ladefoged, P. (1982) *A Course in Phonetics*. San Diego: Harcourt Brace Jovanovich.

Ladefoged, P, Cochran, A. & Disner, S. (1977) Laterals and trills. *Journal of the International Phonetic Association* **7**:46-54.

Lado, R. (1957) *Linguistics Across Cultures*. Ann Arbor: University of Michigan Press.

Lass, R. (1976) *English Phonology and Phonological Theory*. London: Cambridge University Press.

Lehn, W. & Slager, W. (1959) A contrastive study of Egyptian Arabic and American English: The segmental phonemes. *Language Learning* **9**:25–33.

Lenneberg, E. (1967) *Biological Foundations of Language*. New York: Wiley.

Levelt, C. (1994) *On the Acquisition of Place*. Ph.D. Dissertation. Holland Institute of Generative Linguistics, Leiden University.

Levelt, C. (1992) Consonant harmony: A reanalysis in terms of vowel-consonant interaction. Paper presented at the 17th annual Boston University Conference on Language Development: Boston.

Levin, J. (1988) A Place for Lateral in the Feature Geometry. ms., University of Texas, Austin.

Lindau, M. (1985) The story of /r/. In V. Fromkin (ed) *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*, pp. 157-168. New York: Academic Press.

Lombardi, L. (1990) The nonlinear organization of the affricate. *Natural Language and Linguistic Theory* **8**:375-425.

MacKain, K., Best, C. & Strange, W. (1981) Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics* **2**: 369-390.

Macken, M. & Ferguson, C. (1983) Cognitive aspects of phonological development: Model, evidence and issues. In K. Nelson (ed.) *Children's Language*, Vol. 4, pp. 255-282. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Maddieson, I. (1984) *Patterns of Sound*. Cambridge: Cambridge University Press.

Mascaró, J. (1976) *Catalan Phonology and the Phonological Cycle*. Ph.D. dissertation, MIT.

Mattingly, I., Liberman, A., Syrdal, A. & Halwes, T. (1971) Discrimination in speech and non-speech modes. *Cognitive Psychology* **2**:131-57.

McCarthy, J. (1988) Feature geometry and dependency: A review. *Phonetica* **43**:84-104.

McCarthy, J. & Prince, A. (1993) Prosodic Morphology I: Constraint Interaction and Satisfaction. Ms., University of Massachusetts, Amherst, and Rutgers University, New Brunswick, N.J.

McCarthy, J. & Taub, A. (1992) Review of Paradis & Prunet. *Phonology* **9**: 363.

Mehler, J. (1985) Language related dispositions in early infancy. In J. Mehler & R. Fox (eds) *Neonate Cognition: Beyond the Blooming, Buzzing Confusion.* Hillside, NJ: Erlbaum.

Menn, L. (1978) Phonological units in beginning speech. In A. Bell & J. Hooper (eds.) *Syllables and Segments*, pp. 157-171. Amsterdam: North-Holland.

Menn, L. (1977) Phonotactic rules in beginning speech. *Lingua* **26**:225-251.

Mester, A. & Ito, J. (1989) Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language* **65**(2):258-293.

Michaels, D. (1974) Sound replacemenets and phonological systems. *Linguistics* **176**:69-81.

Michaels, D. (1973) Sinhalese sound replacement and feature hierarchies. *Linguistics* **170**:14-22.

Miller, G. & Nicely, P. (1955) An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America* **27**(2). Reprinted in I. Lehiste (ed.) *Readings in Acoustic Phonetics*, pp. 301-315. Cambridge, MA: MIT Press.

Miller, J. & Eimas, P. (1977) Studies on the perception of place and manner of articulation: A comparison of the labial-alveolar and nasal-stop distinctions. *Journal of the Acoustical Society of America* **61**:835-45.

Mithun, M & Basri, H. (1985) The phonology of Selayarese. *Oceanic Linguistics* **25**(1):210-254

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A., Jenkins, J. & Fujimura, O. (1975) An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English. *Perception & Psychophysics* **18**: 331-340.

Munro, M., Flege, J. & MacKay, I. (1996) The effects of age of second language learning on the production of English vowels. *Applied Psycholinguistics* **17**:313-334.

Padgett, J. (1994) Feature classes. In J. Beckman, S. Urbanczyk & L. Walsh (eds.), *Papers in Optimality Theory*, University of Massachusetts Occasional Papers **18**:385-420.

Paradis, C. & Prunet, J.-F. (1989) On coronal transparency. *Phonology* **6**(2):317-348.

Paradis, C. & Prunet, J.-F. (eds.) (1991) *The Special Status of Coronals*. San Diego: Academic Press.

Paradis, C. & LaCharite, D. (eds.) (1993) *Constraint Based Theories in Multilinear Phonology*, The Canadian Journal of Linguistics **38**.

Paradis, M. & Libben, G. (1987) *Assessment of Bilingual Aphasia*. Hillsdale, N.J.: Lawrence Erlbaum.

Pater, J. (1993) Theory and methodology in the study of metrical parameter (re)setting. *McGill Working Papers in Linguistics* **9**: 211-243.

Peterson, G. & Lehiste, I. (1967) Duration of syllable nuclei in English. In I. Lehiste (ed.) *Readings in Acoustic Phonetics*, pp. 191-201. Cambridge: MIT Press.

Piggott, G. (1993) The geometry of sonorant features. ms., McGill University, Montreal, Canada.

Piggott, G. (1992) Variability in Feature Dependency: The Case of Nasality. *Natural Language and Linguistic Theory* **10**:33-77.

Pinker, S. (1984) *Language Learnability and Language Development*. Cambridge, Ma.: Harvard University Press.

Pisoni, D. (1973) Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics* **13**:253-60.

Pisoni, D. (1971) On the Nature of Categorical Perception of Speech Sounds. Unpublished Ph.D. dissertation: The University of Michigan.

Prince, A. & Smolensky, P. (1993) *Optimality Theory: Constraint Interactio in Generative Grammar*, RuCCs Technical Report #2, Rutgers University Center for Cognitive Science, Piscaraway, N.J. [to appear, Cambridge, Mass: MIT Press].

Pulleyblank, D. (1997) Optimality theory and features. In D. Archangeli & D.T. Langendoen (eds.), *Optimality Theory: An Overview*, pp. 59-101. Oxford: Blackwell Publishers.

Pulleyblank, D. (1986) *Tone in Lexical Phonology*. Dordrecht: Reidel.

Pye, C., Ingram, D. & List, H. (1987) A comparison of initial consonant acquisition in English and Quiché. In K. Nelson & A. van Kleek (eds.), *Children's Language*, Vol. 6, pp. 175-190. Hillsdale, NJ: Lawrence Erlbaum Associates.

Radford, A. (1988) Small children's small clauses. *Transactions of the Phonological Society* **86**:1-43.

Raphael, L. (1972) Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *Journal of the Acoustical Society of America* **51**:1296–1303.

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

Renck, G. (1967) A grammar of Yagaria. *Pacific Linguistics* B-40.

Repp, B. (1984) Categorical perception: Issues, methods, findings. In N.J. Lass (eds.) *Speech and Language: Advances in Basic Research and Practice*, Vol. 10. Orlando: Academic Press.

Rice, K. (1996) Default variability: the coronal-velar relationship. *Natural Language and Linguistic Theory* **14**(3):493-543.

Rice, K. (1993) A reexamination of the feature [SONORANT]: The status of 'sonorant obstruents'. *Language* **69**:308-344.

Rice, K. (1992) On deriving sonority: A structural account of sonority relationships. *Phonology* **9**:61-99.

Rice, K. (1989) *A Grammar of Slave*. New York: Mouton de Gruyter.

Rice, K. & Avery, P. (1995) Variability in a deterministic model of language acquisition: A theory of segmental acquisition. In J. Archibald (ed.) *Phonological acquisition and phonological theory*, pp. 23-42. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rice, K. & Avery, P. (1993) Segmental complexity and the structure of inventories. *Toronto Working Papers in Linguistics* **12**(2):131-154.

Rice, K. & Avery, P. (1991a) A Learning Path for Phonology: A Theory of Segmental Elaboration. Paper presented at the 16th Annual Boston University Conference on Language Development. October, 1991.

Rice, K. & Avery, P. (1991b) On the relationship between laterality and coronality. In C. Paradis and J.F. Prunet (eds.) *Phonetics and Phonology: The Special Status of Coronals*, Vol. 2, pp. 101-124. San Diego: Academic Press.

Rice, K. & Avery, P. (1989) On the Interaction Between Sonorancy and Voicing. *Toronto Working Papers in Linguistics* **10**:65-82.

Richmond, S. (1964) *Statistical Analysis*. New York: The Ronald Press Company.

Ritchie, W. (1968) On the explanation of phonic interference. *Language Learning* **18**:183-197.

Sagey, E. (1986) *The Representation of Features and Relations in Non-linear Phonology*, Doctoral dissertation, Massachussetts Institute of Technology, Cambridge, MA.

Sapir, E. & Hoijer, H. (1967) *The Phonology and Morphology of the Navaho Language*. Berekely: University of California Press.

Schachter, J. (1989) Testing a proposed universal. In S. Gass & J. Schachter (eds.) *Linguistic Perspectives on Second Language Acquisition*, pp. 73-88. Cambridge: Cambridge University Press.

Schwartz, B. (in press) On two hypotheses of "transfer" in L2A: Minimal trees and absolute L1 influence. In S. Flynn, G. Martohardjono & W. O'Neil (eds.) *The Generative Study of Second Language Acquisition*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Schwartz, B. (1993) On explicit and negative data effecting and affecting *competence* and *linguistic behavior*. *Studies in Second Language Acquisition* **15**:147-63.

Schwartz, B. & Eubank, L. (eds.) (1996) special issue of*Second Language Research* **12** (1).

Schwartz, B. & Sprouse, R. (1996) L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* **12**(1):40-72.

Scullen, M. (1992) Feature geometry and underspecification in developing phonologies. Paper presented at the 17th annual Boston University Conference on Language Development: Boston.

Shvachkin, N. Kh. (1948) Razvitiye fonematicheskogo vospriyatiya rechi v rannem vozraste. *Izvestiya Akademii Pedagogicheskikh Nauk RSFSR* **13**:101-132. Citations from the English translation, The development of phonemic speech perception in early childhood. Translated by E. Dernbach and republished, 1973, In C. Ferguson & D. Slobin (eds.) *Studies of Child Language Development*, pp. 91-127. New York: Holt, Rinehart and Winston.

Shaw, P. (1991) Consonantal harmony systems: The special status of coronal harmony. In C. Paradis and J.F. Prunet (eds.) *Phonetics and Phonology: The Special Status of Coronals*, pp. 125-157. San Diego: Academic Press.

Sheldon, A. & Strange, W. (1982) The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics* **3**: 243-261.

Smith, N. (1973) *The Acquisition of Phonology*. Cambridge: Cambridge University Press.

Spencer, A. (1986) Towards a theory of phonological development. *Lingua* **68**:3-38.

Spencer, A. (1984) On eliminating the feature [lateral]. *Journal of Linguistics* **20**: 23-43.

Stemberger, J.P. & Stoel-Gammon, C. (1991). The underspecification of coronals: Evidence from language acquisition and performance errors. In C. Paradis and J.F. Prunet (eds.) *Phonetics and Phonology: The Special Status of Coronals*, pp. 181-199. San Diego: Academic Press.

Steriade, D. (1993) Closure, release and nasal contours. In M. Huffman & R. Krakow (eds.) *Phonetics and Phonology*, Vol. 5, pp. 401-470. San Diego: Academic Press.

Steriade, D. (1987) Redundant values. In A. Bosch, B. Need & E. Schiller (eds)*Papers from the CLS Parasession on Autosegmental and Metrical Phonology*, pp. 339-362. Chicago Linguistic Society.

Stevens, K. (1989) On the quantal nature of speech. *Journal of Phonetics* **17**:3-46.

Stockwell, R. & Bowen, J. (1965) *The Sounds of English and Spanish*. Chicago: University of Chicago Press.

Stoel-Gammon, C. & Dunn, C. (1985) *Normal and disordered Phonology in Children*. Baltimore: University Park Press.

Strange, W. & Dittmann, S. (1984) Effects of discrimination training on the perception of /r–l/ by Japanese adults learning English. *Perception & Psychophysics* **36**: 131-145.

Strange, W. (ed.) (1995) *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore: York Press.

Streeter, L. (1976) Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature* **259**: 39-41.

Studdert-Kennedy, M. (1986) Some developments in research on language behavior. In N. Smelser & D. Gerstein (eds.) *Behavioral and social science: Fifty years of discovery*, pp. 208-248. Washington, DC: National Academy Press.

Tees, R. & Werker, J. (1984) Perceptual Flexibility: Maintenance or recovery of the ability to discriminate non-native speech sounds. *Canadian Journal of Psychology* **38**: 579-590.

Trehub, S. (1976) The discrimination of foreign speech contrasts by infants and adults. *Child Development* **47**:466-472.

Trigo, R. (1988) *On the Phonological Derivation and Behavior of Nasal Glides*. Ph.D. dissertation, Massachusetts Institute of Technology.

Trommelen, M. (1984) *The Syllable in Dutch*. Dordrecht: Foris.

Trubetzkoy, N. (1939) *Grundzuge der Phonologie*. Guttingen: Vanderhoeck and Ruprecht. Citations from the English translation, *Principles of Phonology*, translated by C.A.M. Baltaxe, 1969. Berkeley: University of California Press.

Tsushima, T., Takizawa, O., Sasaki, M., Shiraki, S., Nishi, K., Kohno, M., Menyuk, P., & Best, C. (1994) Discrimination of English /r–l/ and /w–y/ by Japanese infants at 6–12 months: Language-specific developmental changes in speech perception abilities. ms., Boston University.

Vance, T. (1987) *An Introduction to Japanese Phonology*. Albany, NY: State University of New York Press.

Vihman, M. (1978) Consonant harmony: Its scope and function in child language. In J. Greenberg (ed.) *Universals of Human Language*, pp. 282-334. Palo Alto: Stanford University Press.

Weinberger, S. (1994) Functional and phonetic constraints on second language phonology. In M. Yavas (ed.) *First and Second Language Phonology*, pp. 283-302. San Diego: Singular Publishing Group.

Weinberger, S. (1990) Minimal segments in second language phonology. In J. Leather and A. James (eds.) *New Sounds 90, Proceedings of the 1990 Amsterdam Symposium on the Acquisition of Second-Language Speech*, pp. . Amsterdam: University of Amsterdam.

Werker, J., Gilbert, J., Humphrey, K. & Tees, R. (1981) Developmental aspects of cross-language speech perception. *Child Development* **52**: 349-353.

Werker, J. & LaLonde, C. (1988) Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology* **24**: 672-683.

Werker, J. & Logan, J. (1985) Cross-language evidence for three factors in speech perception. *Perception and Psychophysics* **37**:35-44.

Werker, J. & Polka, L. (1993) Developmental change in speech perception: new challenges and new directions. *Journal of Phonetics* **21**: 83-101.

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

Werker, J. & Tees, R. (1984b) Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America* **75**: 1866-78.

Wexler, K. & Manzini, R. (1987) Parameters and learnability in binding theory. In T. Roeper & E. Williams (eds.) *Parameter Setting*, pp. 41-76. Dordrecht: Reidel.

White, L. (1989a) *Universal Grammar and Second Language Acquisition*. Amsterdam: John Benjamins Publishing Company.

White, L. (1989b) Linguistic universals, markedness and learnability: Comparing two different approaches. *Second Language Research* **5**(2):127-140.

White, L. (1988) Universal Grammar and language transfer. In J. Pankhurst, M. Sharwood Smith & P. Van Buren (eds.) *Learnability and Second Languages: A Book of Readings*, pp. 36-60. Dordrecht: Foris.

Williams, L. (1977) The perception of stop consonant voicing by Spanish-English bilinguals. *Perception & Psychophysics* **21**:289-97.

Wode, H. (1992) Categorical perception and segmental coding in the ontogeny of sound systems: A universal approach. In C. Ferguson, L. Menn & C. Stoel-Gammon (eds.) *Phonological Development: Models, Research, Implications*, pp. 605-631. Baltimore: York Press.

Wode, H. (1978) The beginnings of L2-phonological acquisition. *IRAL* **16**:109-24.

Wu, Y. (1991) Mandarin suffix [R] revisited. *Proceedings of the 1991 Annual Conference of the Canadian Linguistics Association*. University of Toronto.

Yamada, R. (1995) Age and acquisition of second language speech sounds: Perception of American English /ɹ/ and /l/ by native speakers of Japanese. In W. Strange (ed.) *Speech perception and linguistic experience: Issues in cross-language research*, pp. 305–320. Baltimore: York Press.

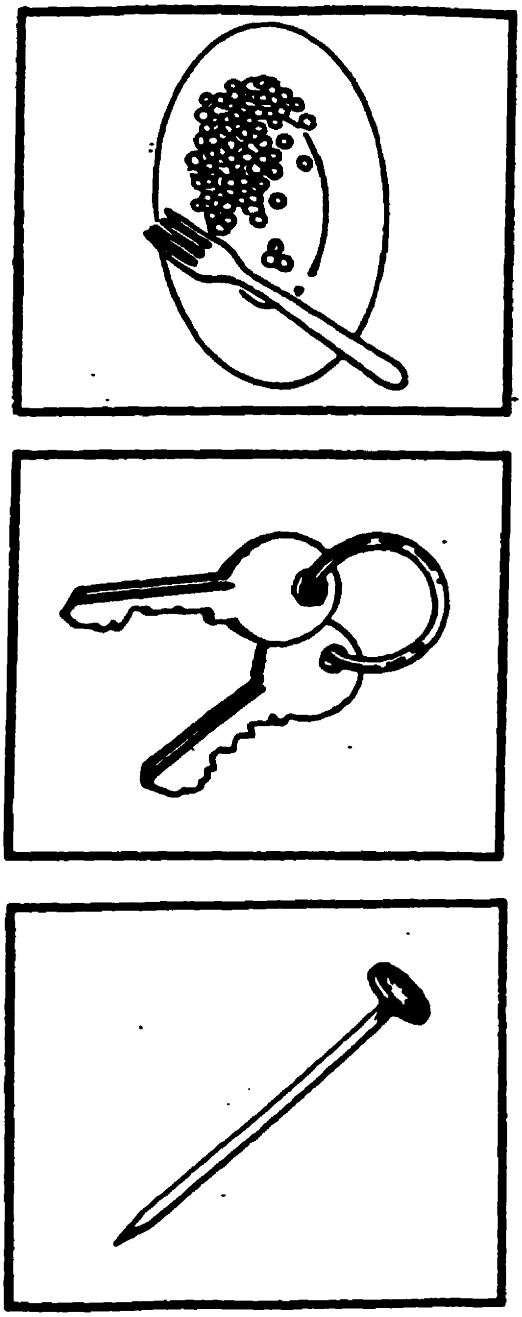
Yamada, R., Kobayashi, N. & Tohkura, Y. (in press) Effect of word familiarity on non-native phoneme perception: Identification of English /r/, /l/ and /w/ by native speakers of Japanese. In A. James & J. Leather (eds.) *Second Language Speech*. New York: Mouton de Gruyter.

Yip, M. (1990) Two Cases of Double-Dependency in Feature Geometry. ms., Brandeis University, Waltham, MA.

Zoll, C. (1993) 'Superclosed vowels' in Bantu: A feature geometric approach. Paper presented at the 24th Annual Conference on African Linguistics. Columbus: The Ohio State University.

Appendix A

Sample stimulus array
(L1 acquisition study)



Appendix B

Experimental stimulus items
(L1 acquisition study)

SV Node
sonorant vs. obstruent

mail	pail
nail	tail
rose	toes
dot	knot
bat	mat

Peripheral Node
coronal vs. labial

pail	tail
map	nap
mail	nail
ball	doll
pie	tie

Oral Node
nasal vs. oral sonorant

light	night
nose	rose
lap	nap
nail	rail
neck	wreck

Peripheral Node
coronal vs. velar

cap	tap
deer	gear
cop	top
key	tea
cape	tape

Vocalic Node
lateral vs. vocalic sonorants

lamb	ram
lock	rock
lake	rake
lamp	ramp
lip	rip

Dorsal Node
labial vs. velar

boat	goat
keys	peas
bait	gate
cot	pot
core	pour

Appendix C

Experimental stimulus items
(L2 acquisition study)

<u>Onset Condition</u>

lip–rip
lock–rock
lice–rice
rake–lake
lamp–ramp
lamb–ram

<u>Cluster Condition</u>

blade–braid
clown–crown
fly–fry
flute–fruit
glass–grass
flock–frock

<u>Coda Condition</u>

ball–bar	[bɒl–bɑr]
eel–ear	[iːl–iːr]
pail–pear	[peɪl–per]
bowl–boar	[boʊl–bor]
tile–tire	[taɪl–taɪr]
hold–hoard	[hoʊld–hord]

<u>Non–l/r Foils</u>

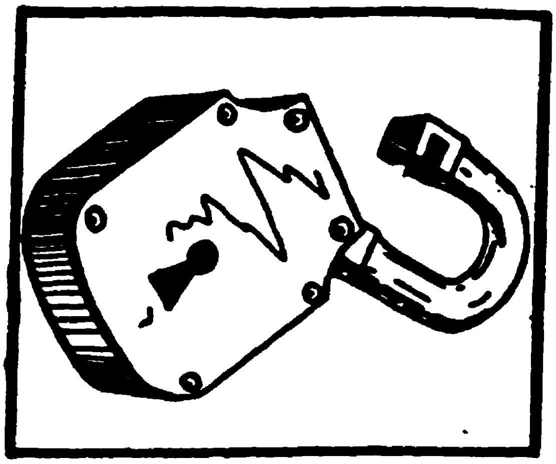
light–night
toes–rose
nose–rose
keys–peas
ball–doll
tap–lap
nose–toes

<u>Identical Pair Foils</u>

lice–lice
rock–rock
lock–lock
fly–fly
fry–fry
braid–braid
blade–blade
pear–pear
pail–pail
hold–hold
hoard–hoard
bowl–bowl
boar–boar
tire–tire
tile–tile

Appendix D

Sample stimulus array
(L2 acquisition study)



Appendix E

An Examination of the Implications of Optimality Theory for First and Second Language Acquisition

The knowledge that speakers possess about their languages does not change despite changes in our formal characterizations of that knowledge; however, our theories of acquisition at a specific point in time *are* dependent upon the particular characterizations of linguistic knowledge at that time. Thus, as the theory of grammar changes, our understanding of how that internal system is acquired may also require modification. Recently, theoretical phonology has witnessed a dramatic shift in the approach to defining grammatical knowledge that has potential consequences for our interpretation (and explanation) of the acquisition and perception data discussed in this thesis. This shift in phonological theory and characterization of linguistic knowledge has been due in large part to the development of Optimality Theory (OT) (Prince & Smolensky, 1993; McCarthy & Prince, 1993).

In this appendix, I consider, fairly closely, the implications that this new theory of grammar has for first and second language acquisition. An important goal of this discussion will be to examine whether the implications of this new conception and characterization of phonological knowledge undermine the theory of first and second language phoneme acquisition developed here; we will also explore the possibility that the experimental findings obtained in this thesis can be brought to bear on current theoretical issues. After elaborating the fundamental premises of OT, I will consider the consequences that adopting each of these premises entails. Experimental evidence from first and second language acquisition will then be brought to bear on these issues. This evidence suggests that some of the consequences that follow from the OT view of the grammar are not supported; I argue, furthermore, that we obtain a more satisfactory explanation of the perception and acquisition data presented in this dissertation under the view of the grammar assumed in Feature Geometry theory.

A.1 A Brief Overview of the Tenets of Optimality Theory

In order to consider the implications of this framework for acquisition and perception, we must first briefly review the main tenets of OT. There are three fundamental premises: 1) the grammar of a language consists of a rank-ordered set of constraints, 2) these constraints are minimally violable, and 3) these constraints are universal, with different languages' grammars being defined by different rankings of the set of constraints. Let us examine each of these in turn.

In standard non-linear phonology, the relationship between an underlying representation and its surface form is captured by the application of a set of rules (in connection with a set of well-formedness constraints) to underlying forms in order to derive surface forms. In this view, the grammar consists of a set of representations, a set of rules and a set of constraints. However, within OT, the grammar is viewed as a set of constraints only; no rules (or derivational processes) are presumed to exist in the grammar, and it is currently being debated to what degree (if at all) representations contain internal structure (e.g., feature-geometric structure). The relationship between underlying representations (termed INPUTS in OT) and their surface forms (termed OUTPUTS) is governed by satisfaction of the ranked set of constraints.

Thus, OT accounts for phonological phenomena through a process of constraint satisfaction: a component of the grammar (termed GEN) generates a set of candidate OUTPUTS which are evaluated according to how well each satisfies the ranked set of constraints; the member of the candidate set that best satisfies the set of constraints in that grammar (i.e., satisfies the highest ranked constraints relevant for that form) will be optimal and hence surface as the OUTPUT form corresponding to that INPUT. Importantly, there are assumed to be no constraints on INPUTS; regardless of the specific content of the INPUT, the effect of the constraint hierarchy is to prevent all but the correct (i.e., optimal) form from being produced. The different rankings of these constraints produce the cross-linguistic variation we observe in phonological data.

Another fundamental premise of this approach is that, in contrast to the strict satisfaction demanded of well-formedness constraints of non-linear phonology, constraints in OT are minimally violable. When one constraint is ranked above another, the requirements of the higher ranked constraint take precedence in determining well-formedness. In such a case, violation of a lower ranked constraint is permissible in order to satisfy a more highly ranked constraint. Thus, an optimal OUTPUT is not one that satisfies all of the well-formedness constraints, but rather the one that violates the fewest number of high-ranking constraints (relative to the other candidates). One of the consequences of minimal violability is that the effect of any given constraint is no longer "all or nothing"; when possible, a constraint will be satisfied, but under certain circumstances it may not be.

The final premise of OT that we will be concerned with here is that the set of constraints is universal (i.e., the same set of constraints is present in every grammar); it is the variable ranking of these constraints that defines the grammars of different languages and produces cross-linguistic variation. In some languages, though, the effects of every constraint may not be readily apparent; this lack of overt visibility of the workings of a particular constraint are due, however, to its relatively low ranking in the grammar rather than its absence from the grammar altogether. Since each constraint is universal and, therefore, present in every grammar, acquisition in OT does not involve the acquisition of language-specific constraints; rather the process of acquisition is the re-ranking of the set of constraints, from some default order given by Universal Grammar (UG) to the specific order for the language being acquired. Thus, in this view, UG consists of the default ranking of the constraints and is itself transmuted into the target language grammar on the basis of positive evidence.[1]

[1] Note that this view is consistent with Chomsky & Lasnik (1993), but differs from Haegman (1991), White (1989a), and the view taken in this dissertation that properties of UG are transmitted to the emerging grammar, but that UG itself remains intact.

The notion that different grammars are defined by different constraint rankings and the effects of minimal violability and can be seen in the generic tableaux for two hypothetical languages in (1). Both languages contain the same constraints (here Constraint A, B and C), but they differ with respect to the rank-ordering of those constraints.

(1)

Language X	Constraint A	Constraint B	Constraint C
/INPUT/			
☞ candidate 1		*	
candidate 2	*!		

Language Y	Constraint B	Constraint A	Constraint C
/INPUT/			
candidate 1	*!		
☞ candidate 2		*	

* = Constraint violation
*! = Fatal violation (candidate no longer considered)
☞ = Optimal candidate
= Constraint not relevant for determining the well-formedness of this candidate

Assuming the same INPUT, GEN will generate the same set of candidate OUTPUTS in both languages (only two are listed since they suffice to illustrate the point here); however, the different rankings will obtain different optimal OUTPUTS (candidate 1 in language X, candidate 2 in language Y). Thus, the cross-language differences illustrated here between these two languages are a result of the reverse ranking of Constraints A and B.

The notion of minimal violability is also illustrated by the above tableaux. Although candidates 1 and 2 are the optimal OUTPUTS in Languages X and Y, respectively, these forms nevertheless violate constraints in the language. However, the optimal OUTPUT violates the least number of more highly ranked constraints. Thus, for example, candidate 1 in Language X is the optimal OUTPUT because it does not violate the highest ranked constraint here, Constraint A, eventhough it does violate the more lowly ranked constraint, Constraint B.

A.2 Some Consequences of these Premises

This section outlines some of the consequences that follow from these fundamental premises. There are two issues, in particular, that I would like to focus on: 1) the consequences that the emphasis on constraints in OT has for our theory of segment structure, and 2) the consequences that the notion of constraint violability has for our theory of underspecification. Each of these points will be elaborated, paying close attention to the implications that each has for acquisition, followed by a discussion of whether the experimental data discussed in previous chapters can be used to resolve some of these issues.

The Internal Structure of Segments

Concomitant with the move in OT to account for phonological phenomena in terms of constrain satisfaction, there has been a shift away from an emphasis on segmental-internal structure (e.g., Feature Geometry theory). If we assume that the relationship between INPUT and OUTPUT forms is governed solely by a set of constraints, there may be little need to posit highly articulated representations. Thus, OT de-emphasizes (highly-articulated) representations, placing the burden of accounting for phonological alterations instead onto constraints and their interaction.

There have been several recent attempts to abolish feature-geometric structure within the segment and to derive the hierarchical organization of features from the ranking and interaction of constraints on the occurrence of features (e.g., Cole & Kisseberth, 1994; Padgett, 1994; Pulleyblank, 1997; cf. esp. Kawasaki, 1997). Since this approach runs counter to the theories of first and second language phoneme acquisition developed in this thesis, it is important to consider the extent to which the OT framework can adequately account for the properties of acquisition and speech perception that were argued in this thesis to follow from the hierarchical organization of features within the segment.

Before examining how OT might capture the relations that seem to hold between certain features, let us (re)consider what specific properties of phoneme acquisition such an OT analysis would have to explain. We saw that children acquire phonemic contrasts in a uniform order. The dependency relations encoded in the feature hierarchy enable us to explain this fact: structurally less complex segments are acquired before structurally more complex segments (i.e., superordinate structure is acquired before dependent structure); relations between particular features predict the specific orders of acquisition that we observe (e.g., labial segments are distinguished from non-labial segments before a contrast is made between the non-labials).

Furthermore, constituency relations within the universal feature hierarchy enable us to explain the limited amount of variability observed across children. The Organizing nodes SONORANT VOICE and PLACE, which dominate the features pertaining to sonorancy (e.g., [approximant], [nasal]) and place of articulation (e.g., [labial], [dorsal]), respectively, are independent constituents and may, therefore, be acquired independently. The order of elaboration *within* a single Organizing node is fixed, but the order of elaboration *across* the various Organizing nodes is not restricted. Given the independence of these Organizing nodes, some children will first acquire contrasts of sonorancy, some children will first acquire contrasts within the place of articulation, and some children will elaborate both Organizing nodes simultaneously. Unless we assume hierarchical relations between features (encoded *somewhere*), we lose an insightful explanation for the uniformity across children (including the limited amount of variability) and for the specific orders of phoneme acquisition that are observed.

Our account of the L1 grammar's influence on speech perception and L2 phoneme acquisition also relies on the hierarchical relation of features. It is, for example, the dependence of [anterior] and [retroflex] on the Coronal node that enables us to explain why speakers whose L1 grammars lack those dependent features fail to discriminate contrasts within the coronal place of articulation (i.e., anterior coronals from retroflex coronals), perceiving all such sounds simply as coronal. Without this sort of dependency relation (again, encoded *somewhere*), we would be unable to predict which L1 phonemic category a given L2 segment will be perceived as. Thus, the dependencies between features is clearly an important factor in L1 and L2 phoneme acquisition; these theoretical notions appear to be a real component of speakers' knowledge of phonemes. The following section outlines how the relations between features can be encoded into constraints (and their rankings); we will see that there may be little empirical difference between OT and Feature Geometry theory with respect to acquisition.

There are two types of constraints posited by OT that are relevant for our discussion: faithfulness constraints and *Structure constraints. Faithfulness constraints serve to maintain identity between the INPUT and the OUTPUT. That is, faithfulness constraints ensure that the phonological material that is present in the INPUT will also be present in the OUTPUT. An example of this kind of constraint is PARSE F (where F stands for some phonological feature); satisfaction of PARSE F assures that a feature present in the INPUT will be parsed and surface as an element of the OUTPUT. The other type of constraint, *Structure, favors OUTPUTS with the fewest features possible; thus, in contrast to faithfulness constraints, *Structure constraints typically cause an OUTPUT to diverge from strict identity with the INPUT. For example, satisfaction of *F will ensure that that feature is not present in the OUTPUT, regardless of whether it is present in the INPUT.

The relative ranking of PARSE F over *F in a given constraint hierarchy (i.e., grammar) encodes the fact that the feature F is contrastive in that particular language: the presence of the feature F in an INPUT will be maintained in the OUTPUT because the satisfaction of PARSE F is of greater importance in this language than the restriction against that feature appearing in the OUTPUT, *F. Each phonological feature will have a corresponding faithfulness constraint and *Structure constraint that in tandem determine its appearance in OUTPUT forms; for example, there will be a PARSE [Coronal] and a *[Coronal], a PARSE [aspiration] and a *[aspiration], and so on.

One of the fundamental premises of OT is that all potential rankings of the universal set of constraints are possible, with each different ranking yielding a possible human language. With respect to the various PARSE F and *F constraints (i.e., one of each for each feature), this means that any rank ordering of these with respect to one another is possible, and should yield a licit phoneme inventory. In other words, (in theory) there is no inherent ordering between any of the constraints that govern the appearance of features in OUTPUT forms: the inclusion of one feature in a language, for example [anterior] (i.e., PARSE [anterior] >> *[anterior]), does not entail the presence of any other particular feature, for example [Coronal] (i.e., PARSE [Coronal] >> *[Coronal]).[2]

It has come to be realized though, that certain features have inherent relations between each other; the typical case appeals to phonetic motivation, in which the movement of one articulator delimits the corresponding constriction location (e.g., the feature [Coronal] delimits constriction to the region in front of the soft palate). In order to capture this sort of inherent relation (as well as other non-segmental phenomena), OT has had to introduce the idea that the rankings of some constraints with respect to each other are fixed. This so- called "harmonic ranking" of constraints maintains a strict order within a small subset of constraints; however, these constraints may be interdigitated with other constraints. It is only the rank ordering of the harmonically-ranked constraints with respect to one another that is universally invariant. Thus, the notion of harmonic ranking provides a way for (at least some of) the "dependencies" between features, captured structurally in Feature Geometry, to be derived using constraint ranking in OT.

Let us now consider whether the L1 phoneme acquisition data reported in Chapter 2 can be accounted for under the assumption that the inherent relations between features can be derived by the fixed rankings of constraints. The first aspect of the L1 data that must be explained is the fact that the acquisition of phonemic contrasts is

---

[2] Compare this position with Feature Geometry theory in which the presence of a dependent, in this example [anterior], forces the presence of that feature's superordinate feature, here [Coronal].

gradual; young children begin unable to distinguish segments phonologically (i.e., in their grammars) and then one by one phonemic contrasts are acquired. Since it is the ranking of PARSE F >> *F that entails that that feature contributes to defining the inventory of a language, then we must assume that acquisition of phonemic contrasts in OT would involve rank ordering PARSE F over *F. Recall that in OT, all languages contain all constraints. Thus, Universal Grammar is believed to contain the full set of constraints; the child's acquisition task is to determine the correct rank- ordering of those constraints for the target language. In this view, UG (the default ranking of constraints) is itself altered, on the basis of positive evidence, into the target language. If we assume, along the lines of Gnanadesikan (1995), that the default ranking of constraints given by UG is *F >> PARSE F, then when the child detects that a given feature is used contrastively (i.e., distinguishes phonemes) in the ambient language, PARSE F will be re-ranked over *F.[3] Therefore, from this brief examination, it appears that the *gradual* acquisition of phonemic contrasts on the basis of positive evidence can be explained with OT.

A potentially more difficult aspect of the L1 data to explain in terms of OT constraint rankings is the uniform order of acquisition across English children, as well as the specific order of acquisition within each child. Yet, here too, OT may be able to account for the developmental facts. By assuming that the PARSE constraints for various features are harmonically ranked with respect to one another (and the corresponding *Structure constraints are also harmonically ranked with respect to one another), we can ensure that the order in which the different features are introduced into a language (i.e., added to the list of features that defines the phoneme inventory) is restricted. The promotion of a PARSE constraint low in the fixed ranking would entail the concomitant promotion of those harmonically-ranked PARSE constraints above it. For example, if we assume on learnability grounds that faithfulness constraints are

---

[3] Note that since the ranking of PARSE F over *F indicates a contrast in the grammr, the opposite default ranking, PARSE F>>*F, would entail that children *from the beginning* would phonologically distinguish all segments. We know this to be false, however, from Chapter 1. Furthermore, this default ranking would require negative evidence of some sort to re-rank (e.g., the child would have to be aware of lack of a contrast in the ambient language), either to *promote* *F or to *demote* PARSE F.

initially more highly ranked than the *Structure contrasts, then the fixed ranking of PARSE[Labial] >> PARSE[Dorsal] >> PARSE[Coronal] would mean that the feature [Labial] could be independently "added" to the phoneme inventory (as we indeed found in the acquisition data), while the acquistion (i.e., promotion) of [Dorsal] would entail acquisition of [Labial], and the acquistion of [Coronal] would entail the acquisition of both [Labial] and [Dorsal].[4]

The variability that we observed between the acquisition of sonorancy and place features could also be captured in terms of harmonic ranking: while the features defining sonorancy (e.g., [nasal], [approximant], etc.) and the place of articulation features (e.g., [Coronal], [anterior], [retroflex], [labial], etc.) could have a fixed rank-ordering amongst themselves, there would be no fixed order between the two groups of constraints with respect to one another. Hence, the promotion of PARSE F constraints could take place within each harmonic ranking independently.

It is clear that the harmonic ranking of constraints is absolutely crucial in order to account for the L1 phoneme acquisition data. Without a fixed ranking, children could freely promote any PARSE F constraint over any other, thereby introducing features to their inventories in an unconstrained fashion; in this scenario, we would expect to find no uniformity, nor any particular order, in the acquisition of phonemic contrasts across the learners of a given language.[5] We know this to be empirically false. We would also expect no uniformity across learners of different languages. On this point, OT and Feature Geometry theory make very different predictions: the relations encoded in Feature Geometry theory are universal; thus similar acquisition orders are predicted for

[4] Note that we must also assume that corresponding *Structure constraints are harmonically ranked in the reverse order: *[Coronal] >> *[Dorsal] >> *[Labial]. This assumption is somewhat intuitive if we view constraints as markedness statements themselves; that is, if the unmarked status of a feature is captured in the ranking of constraints, then the constraint that prevents its appearance (e.g., *[Coronal]) should be high relative to *Structure constraints on the other features, but the constraint that ensures its appearance (e.g., PARSE[Coronal]) should be relatively low with respect to the other faithfulness constraints.

[5] On this point, one might argue that properties of the input (e.g., frequency of occurrence of various phonemes) will result in more or less uniform acquisition orders, without need to resort to the harmonic ranking of featural constraints. These two possible explanations could be tested with an examination of acquisition orders cross-linguistically: the "input" hypothesis would predict no necessary uniformity cross-linguistically, given variable phoneme frequencies in the input, whereas harmonic ranking (like Feature Geometry theory) predicts a substantial amount of cross-linguistic uniformity in the order of phoneme acquisition.

children learning a given language and cross-linguistically. Clearly cross-linguistic data are required to complete our picture of phoneme acquisition.

At this point, then, there appears to be no empirical difference between harmonic ranking of constraints within OT and Feature Geometry theory. Both are able to account for the gradual acquisition of phonemic contrasts and, assuming that the constraints governing the occurrence of features may be ranked in a universally fixed order, both are able to explain the restricted orders that are observed. Thus, the adoption of OT does not appear to undermine the theory of phoneme acquisition outlined in this dissertation; however, the acquisition facts do suggest that the harmonic ranking of featural constraints is required in OT. Thus, as Feature Geometry and harmonically-ranked OT appear to be little more than notational variants, there is no clear advantage in adopting the position that segments have no internal structure. I will continue, therefore, to assume in the remaining discussion that features are organized hierarchically within the segment as a principle of UG (i.e., as a condition on both the INPUT and the OUTPUT). We turn now to the theory of underspecification to consider whether our acquisition and perception data reveal any empirical differences between the two theories with respect to feature specification.

Underspecification

Another component central to our feature-geometric account of L1 and L2 phoneme acquisition is a theory of underspecification: the assumption that features which do not differentiate segments within an inventory are not present in the underlying representation of segments in that language. While not a property of Feature Geometry *per se*, underspecification has played a vital role in determining the correct representation of individual segments in a grammar. By underspecifying certain features in underlying representations, we formally capture properties of the phonological system (e.g., that these redundant features are inert and do not participate in phonological processes). Differences between languages, then, are determined by examining their *underlying* (i.e., distinctive) features; the crucial point here is that, in Feature Geometry theory, cross-language variation in phoneme inventories and phonological processes (as well as acquisition and non-native speech perception, I would argue) follow from the underlying representations. Thus, underspecification also has consequences for our theory of how these representations are acquired and how these representation operate in the perception of speech sounds. The formal distinction between redundant and non-redundant phonological information predicts that these two types of features will have differing involvement in the acquisition of segmental representations and that they will have different psycholinguistic effects once acquired.

The following section outlines how underspecification is derived through constraint satisfaction in OT, and demonstrates that the means of capturing underspecification through constraints currently available within OT run into several logical problems when one tries to account for the observed perception data.

To begin, let us take a brief look at what aspects of L1 acquisition and L1 phonological interference a theory of underspecification enable us to explain. The theory of underspecification, coupled with the universal hierarchy, determines which features the child must acquire, or in other words, how far the feature geometry must be elaborated until the adult geometry is attained. Underspecification also plays a crucial role in our understanding of L1 phonological interference: it is only by assuming that certain features are not present in underlying forms (and, hence, not present in the adult feature geometry) that we are able to explain why those features do not impinge upon speech perception. To take a specific example, though Japanese and Chinese both contain coronal sounds in their phoneme inventories, the theory of underspecification entails that the feature [Coronal] will not be utilized in the phonology of Japanese (because no phonemes are distinguished by this feature, in contrast to Chinese); therefore [Coronal] will be absent from the Japanese L1 grammar and absence of this feature prevents Japanese speakers from accurately perceiving L2 sounds which are distinguished by this feature. Thus, it is only by assuming featural underspecification that we arrive at a principled theory of interference that explains the differential perceptual capabilities of Japanese and Chinese speakers.

Let us turn now to OT. Recall that one of the fundamental tenets of this theory is that constraints are minimally violable. One of the consequences of this premise is that the effects of constraints are no longer "all or nothing"; with respect to features, this means that it is no longer the case that, at any given "level" of the phonology, the features of a segment will be underspecified or fully specified. Instead, underspecification is seen as a property derived from constraint interaction. As a result, OUTPUTS may show underspecification for some set of segments in some subset of environments (driven by the demands of particular highly ranked constraints). A consequence of this is that a given feature is longer seen to be "underspecified" in the underlying representation. In fact, demanding underspecification on this "level" of representation is inconsistent with the OT premise that there are no constraints on INPUTS. A feature's presence or absence in the optimal OUTPUT candidate is determined solely by the pressures of satisfying the most highly-ranked relevant constraint. Thus, in OT one cannot maintain that Feature F is always active or always inactive in a given language; what results is a type of "partial" underspecification (see e.g., Itô, Mester & Padgett, 1995).

Since there are no constraints on INPUT forms, the underlying (i.e., INPUT) forms no longer serve to contrast languages.[6] Regardless of what one assumes to be the INPUT (e.g., whether it is specified for the feature [Coronal] or not), the appropriate ranking of constraints will obtain the surface facts (e.g., whether [Coronal] is present in OUTPUT forms; in other words, whether the feature is contrastive in the language or not). One consequence of this is that we can reasonably assume identical INPUT forms for both Chinese and Japanese (at least with respect to feature specification). There is,

[6] OT researchers are careful to include the following caveat to the claim that there are no constraints on INPUTS: "Inputs may only consist of legitimate phonological material.....and all such material must obey certain basic and universally respected properties of phonological well-formedness. For example, coronality (use of the tongue tip or blade in an articulaton) cannot be represented as a feature of the lower lip in some language, either underlyingly or on the surface. The hypothesis of 'no constraints on input forms' is that none of the *violable* constraints can be imposed as a requirement of inputs..." (Pulleyblank, 1997:78).

nevertheless, clearly a difference between the knowledge that Japanese and Chinese speakers have about their respective phoneme inventories and how that knowledge constrains their perceptual capabilities. I have argued in this dissertation that the difference in the Japanese and Chinese speakers' ability to perceive various English contrasts is a direct result of the underlying featural properties of their grammars. However, if the INPUT forms cannot distinguish the L1 knowledge that bears on each group's perception of speech sounds, where in OT can we locate the relevant difference?

Languages will be distinguished from one another both by their rankings of constraints and by the OUTPUT forms that are produced through satisfaction of those ranked constraints. Thus, there are two types of knowledge that could potentially serve to distinguish the perceptual abilities of speakers of different languages: the OUTPUT forms and the constraints (and their rankings) that produce those OUTPUTS. Underspecification would appear to be a property in OT of both the OUTPUT (i.e., whether certain features are contained in the optimal OUTPUT) and the constraint rankings (i.e., arise from the particular ranking that determines whether that feature is contained in the optimal OUTPUT).

We will first consider the "underspecification" of features in OUTPUTS to determine whether this level of linguistic knowledge (i.e., the surface forms) impinges on perception: do the OUTPUT forms in Japanese and Chinese accurately predict a difference in the ability to perceive English contrasts among speakers of these two languages? The OUTPUTs given in (2) correspond to the INPUT /s/ in Japanese and Chinese, respectively (only the material that distinguishes these two segments is included). How these OUTPUTS are obtained through constraint satisfaction will be outlined below; the focus here is on the OUTPUTS themselves.

(2)

Japanese	*Chinese*
[s]	[s]
ROOT	ROOT
|	|
PLACE	PLACE
	|
	Coronal

Here we see that the OUTPUT forms do indeed distinguish Japanese and Chinese. Hence, this OUTPUT (i.e., surface) level of knowledge (rather than our underlying forms in Feature Geometry) offers a potential explanation for the observed perception facts: under this analysis, the presence of [Coronal] in the Chinese OUTPUT will permit accurate perception of coronal contrasts, whereas its absence from the Japanese OUTPUT will inhibit this ability. It would seem, then, that underlying distinctions between languages' feature inventories are not needed in order to account for the L2 data discussed in this dissertation; underspecification would appear to be a property of OUTPUTS (i.e., in those cases where the satisfaction of constraints prevents featural material present in the INPUT from being realized in the optimal OUTPUT), correctly distinguishing the knowledge different speakers possess about their respective L1s.

Another comparison of OUTPUTS however, this time from English and Thai, shows that OUTPUTS do not reliably distinguish speakers' knowledge of phonemes (and the features that distinguish those phonemes). Both English and Thai contain aspirated voiceless obstruents in the onset position of a syllable. Although aspiration in Thai is contrastive (i.e., aspirated and unaspirated stops are distinct phonemes; this contrast is neutralized in coda position) and not contrastive in English (aspirated voiceless obstruents are in complementary distribution with unaspirated voiceless obstruents), both languages will have aspiration in their OUTPUTS in syllable-initial position. Two such examples are given in (3); again, only the relevant material is included (LAR and asp stand for LARYNGEAL and aspiration, respectively).

(3) *English* (onset) | *Thai* (onset)

English (onset)		Thai (onset)	
[p^h]		[p^h]	
ROOT		ROOT	
LAR	PLACE	LAR	PLACE
asp	Labial	asp	Labial

Keep in mind that since these are OUTPUTS, they represent what is actually produced; although English and Thai differ with respect to whether aspiration is contrastive, they both contain aspirated voiceless stops. Following the logic laid out in the previous paragraph, if presence (or absence) of a feature in the OUTPUT determines accurate perception of non-native contrasts that differ along that dimension, then we would predict that English and Thai speakers should accurately perceive distinctions between aspirated and unaspirated stops. The fact, though, is that English speakers have been shown to be unable to phonologically discriminate segments that differ only in terms of aspiration (Curtin, Goad & Pater, to appear).

Thus, the OUTPUTS of a grammar do not adequately account for cross-language perception patterns. Despite the fact that segments may be "underspecified" in the OUTPUTS due to the ranking and interaction of different constraints, we cannot use this level of knowledge to explain the effect of the L1 phonological system (i.e., grammatical knowledge – which throughout this thesis has been shown yield important insights) on the speech perception system.

The other way in which underspecification is encoded in OT is in the constraints themselves. This level of knowledge could also potentially serve to distinguish the perception abilities of the Japanese and Chinese speakers. Hence, we need to consider whether the variable rankings of specific constraints that produced the OUTPUTS in examples (2) and (3) above can obtain the appropriate cross-language perception differences; if the rank ordering of the constraints themselves cannot adequately encode the notion of underspecification crucial to our account of L2 phoneme acquisition, then we shall be forced to conclude that the move toward eliminating underspecification as a constraint on INPUTS is not empirically supported. What follows is a brief examination of the relevant constraints in Japanese, Chinese, English and Thai.

Recall that the relative ranking of PARSE F over *F in a given constraint hierarchy encodes the fact that the feature F is contrastive in that particular language. Thus, the ranking of PARSE [Coronal] over *[Coronal] would yield a language like Mandarin Chinese in which the feature [Coronal] contributes to defining its phoneme inventory; the reverse ranking of *[Coronal] over PARSE [Coronal] would yield a language like Japanese where Coronal does not serve to distinguish phonemes and is not active in the phonology. Thus, "underspecification" of a given feature (in the OUTPUT) is obtained by the lower ranking of *F in relation to PARSE F.

The two tableaux below in (4) illustrate the ranking of the relevant faithfulness and *Structure constraints in Japanese and Chinese and demonstrate how the satisfaction of these constraints yields the OUTPUTS given above in (2). Note that for the sake of this discussion, I will assume full specification of features in the INPUT; however, as there are no constraints on INPUTS, the optimal OUTPUTS must be obtained regardless of the degree to which those features are specified in INPUTS; also, the dotted line between constraints indicates that they are not ranked with respect to each other (i.e., neither dominates the other), while the solid line indicates that the ranking between the two constraints separated by the line is crucial.

The relative ranking in the Japanese grammar of the constraints PARSE[retroflex], *retroflex, PARSE [Coronal], *[Coronal], reflects the fact that neither the feature [Coronal] nor [retroflex] is contrastive in this language.[7] The INPUT corresponding to the surface form /sa/ may (in theory) be specified for any range of features, including [Coronal] and [retroflex]. GEN then generates a set of candidates that will be evaluated in accordance with the constraints; only two members of that large set are listed in this example, as they are sufficient to establish the relative ranking of the PARSE and *Structure constraints we are interested in.

[7] Of course, this is very small subset of the featural constraints (let alone prosodic constraints) that will comprise any given grammar. Only those constraints that govern the distribution of [Coronal] and [retroflex] are discussed here.

(4)

Japanese	*retro	*Cor	PARSE (retro)	PARSE (Cor)
/sa/ – R – PL – [cor]				
☞ sa – R – PL				
sa – R – PL – [cor]		*!		

Chinese	PARSE (retro)	PARSE (Cor)	*retro	*Cor
/sa/ – R – PL – [cor]				
sa – R – PL		*!		
☞ sa – R – PL – [cor]				

Japanese	*retro	*Cor	PARSE (retro)	PARSE (Cor)
/ṣa/ – R – PL – [cor] – [retro]				
☞ ṣa – R – PL				
ṣa – R – PL – [cor]		*!		
ṣa – R – PL – [cor] – [retro]	*!	*		

Chinese	PARSE (retro)	PARSE (Cor)	*retro	*Cor
/ṣa/ – R – PL – [cor] – [retro]				
ṣa – R – PL	*!	*		
ṣa – R – PL – [cor]	*!			
☞ ṣa – R – PL – [cor] – [retro]				

The important point to notice is that the ranking of *Coronal over PARSE [Coronal] ensures that the features [Coronal] and [retroflex], regardless of whether they are present in the INPUT, are never maintained in the OUTPUT. Thus, not only does this constraint ranking capture the fact that there is no contrast within the class of coronals (e.g., [s] vs. [ṣ]), it also entails that [Coronal] plays no role in the phonology of Japanese; it could potentially also explain why [Coronal] does not play a role in speech perception.

Looking at the tableau for Chinese, we see the opposite ranking: PARSE [Coronal] is ranked higher than *Coronal, and PARSE (retroflex) is ranked higher than *retroflex. This ranking ensures that whenever these two features are part of the INPUT they will be maintained in the OUTPUT. Although we assumed the same INPUTS for Japanese and Chinese, only the Chinese grammar produced a surface contrast between coronal fricatives. Thus, once again we find a difference in the grammar of the Chinese and Japanese speakers, this time in the relative ranking of PARSE [Coronal] and *Coronal: the Chinese grammar's ranking is PARSE [Coronal] >> *Coronal, while the ranking in the Japanese grammar is *Coronal >> PARSE [Coronal]. We might hypothesize, therefore, that the underspecification which is a consequence of constraint rankings is responsible for the differential psycholinguistic effects we have measured.

Examination of the English and Thai rank orderings of the pertinent constraints confirms this hypothesis. Recall that in English, aspiration is allophonic, whereas it is contrastive in onset position in Thai. This difference between languages that use a feature contrastively (even if that feature is neutralized in some contexts) and those in which the feature surfaces in some contexts but does not define phonemes, has been analyzed by Pulleyblank (1997) as resulting from the interaction between PARSE F, *F, and a third constraint that forces the occurrence of that feature in certain circumstances whether or not the feature contributes to defining the inventory. For this discussion, the relevant third constraint is "$_{\text{onset}}$[voiceless/aspirated", captures the fact that cross-linguistically voiceless stops are often realized with delayed voice onset time in onsets; this constraint is presumed to have phonetic motivation, for example, that voicing contrasts in stops are enhanced by aspiration of one member of the contrast.[8] . The interaction of these three constraints is shown in English and Thai is illustrated in (5).[9]

---

[8] The complementary distribution of aspiration in English, and the neutralization of laryngeal contrasts in Thai in the coda position of a syllable, will be obtained by an additional constraint (e.g., Contrastive Coda, Pulleyblank, 1997) which will force the loss of contrasts in coda position. Since the interaction of this additional constraint does not affect the argument here, it is omitted from the tableau and our discussion.

[9] Whether or not /t/ has a feature such as [-voice] in the input does not affect my argument.

(5)

English	Vl/asp [onset]	*asp	PARSE (asp)
/ta/			
ta R L PL	*!		
☞ t^ha R L PL [asp]			
/t^ha/			
ta R L PL	*!		
☞ t^ha R L PL [asp]			
/at^h/[10]			
☞ at R L PL			*
at^h R L PL [asp]		*!	

Thai	PARSE (asp)	*asp	Vl/asp
/ta/			
☞ ta R L PL			
t^ha R L PL [asp]		*!	
/t^ha/			
ta R L PL	*!		
☞ t^ha R L PL [asp]			

The ranking of *asp over PARSE (asp) captures the fact that aspiration is not contrastive in English (i.e., the constraints never derive a surface distinction between voiceless aspirated stops and voiceless unaspirated stops in the same syllabic position); the ranking of Vl/asp over *asp entails that, although aspiration is not contrastive, it will appear in the OUTPUT in well-defined contexts. With respect to PARSE F and *F, English resembles Japanese, in that *F is ranked higher than PARSE F.

Thai, on the other hand, resembles Chinese in its relative ranking of the relevant faithfulness and *Structure constraints, as PARSE (asp) is ranked above *asp. This

[10] This INPUT has been included in order to determine the relative ranking of *asp and PARSE(asp); to do this required an example that would have aspiration in the INPUT that did not appear in the OUTPUT (i.e., not in onset position). This example is not included in the tableau for Thai, however, as the first two examples are sufficient to rank the relevant constraints.

ranking introduces a surface contrast between segments like [p] and [p^h], as illustrated in the tableau. Note that the low ranking of Vl/asp with respect to PARSE (asp), in contrast to English, ensures that any aspiration not present in the INPUT, but introduced by GEN, will not be present in the optimal OUTPUT.

Recall the hypothesis suggested by the comparison of the Japanese and Chinese constraint rankings: the ranking of PARSE F over *F in a speaker's grammar permits him or her to perceive distinctions along the dimension defined by that feature (the case of the Chinese speakers); or, conversely, the ranking of *F over PARSE F prevents a speaker from detecting contrasts that are defined by that feature (the case of the Japanese speakers). Since PARSE (asp) is more highly ranked than *asp in the grammar of Thai speakers, under the constraint analysis, we would predict that Thai speakers would accurately perceive aspiration contrasts in an L2; in contrast, English speakers, with *asp ranked above PARSE (asp), should be unable to phonologically discriminate aspirated and unaspirated stops in an L2. This prediction is indeed borne out by experimental data. Thus, it appears that one can correctly account for cross-language perception and acquisition differences by encoding underspecification of features directly in the constraints; this suggests that the acquisition data support the recent move toward eliminating underspecification as a constraint on INPUTS.

However, if we look closely at the implications that this position makes for the development of perception abilities and phonological categories by children, as well as for the perception and acquisition of non-native contrasts, we will see that there are serious logical problems with the constraint hypothesis; moreover, these problems are not easily surmountable within OT, unless we assume that underspecification is a property of underlying (i.e., INPUT) forms. We will consider each ranking in turn.

It makes no difference whether we posit that the ranking PARSE F >> *F in the adult grammar permits accurate perception of a contrast distinguished by that feature or that ranking *F >> PARSE F inhibits accurate perception; both encounter serious logical and empirical problems when we examine how such a situation might originate. Since PARSE F over *F enables perception and infants are able to hear the vast range of phonetic contrasts employed across the world's languages, we might hypothesize that PARSE F >> *F is the default ranking of these constraints in UG. However, this assumption encounters two serious problems. Recall that it is also this ranking of constraints that designates the contrastive use of that feature in the phonology. Hence, this position would predict that children would phonologically distinguish all of the ambient language's segments from the very beginning We know from the acquisition data reported in Chapter 2, though, that this is false; children gradually acquire the ability to phonologically differentiate segments in their grammar.

Moreover, the default ranking of PARSE F over *F would make it very difficult (if not impossible) to arrive at the appropriate phonemic inventory; no language makes use of all the possible phonemic contrasts, so the specific inventory derived from the constraint rankings would have to be reduced. This reduction would require the detection of the *lack* of contrastive use of a given feature in order to demote the pertinent PARSE F to obtain the *F >> PARSE F ranking. It is fairly certain that children are not sensitive to the type of negative evidence that would be required to recover from a default setting of PARSE F over *F.

Let us then postulate that *F over PARSE F is the default ranking. This ranking would sidestep the negative evidence problem mentioned above (see e.g., Gnanadesikan, 1995); however, it too suffers from serious problems. Since we have hypothesized that *F ranked over PARSE F prevents accurate perception of non-native contrasts (e.g., the case of the Japanese speakers), the *F >> PARSE F ranking should prevent infants from perceiving any contrasts in the input, which we know, of course, to be wrong. Moreover, if the *F >> PARSE F ranking impedes perception, and this were the default, then children would never be able to detect a contrast in the input, in order to promote PARSE F and, thereby, introduce a phonemic contrast into their grammar. Thus, this default ranking also encounters several logical problems and is simply not consistent with the empirical data we have.

It might be instructive to step back for a moment to consider exactly why both formulations of the "constraint" hypothesis suffer from logical inadequacies and fail to account for the developmental data. In the adult speaker, the perception facts and phonemic contrasts in the inventory go hand in hand: both can be explained by a single ranking of PARSE F over *F. However, these two abilities do not converge in infants or young children and cannot, therefore, be captured by a single constraint ranking; they are able to perceive auditorily far more contrasts (i.e., PARSE F over *F) than they are able to differentiate phonologically (i.e., *F over PARSE F).

Given this situation, the position we would like to take is that perception abilities are initially independent of any linguistic knowledge, but that as the L1 grammar is acquired it impinges upon speech perception in such a way that utimately (i.e., in the mature speaker) there is a convergence of perceptual and phonological abilities. This position is not problematic for Feature Geometry theory, as I have demonstrated throughout this thesis. However, it is a significant problem for OT, given its conception of UG's relation to L1 grammar.

In Feature Geometry theory we are able to postulate that structure is transmitted from UG to the emerging grammar, providing for a period (prior to the acquisition of phonological structure) in which speech perception is not encumbered by the grammar, and also allowing for the possibility that phonological knowledge may impinge on the speech perception system, once acquired. In OT, however, UG *is* the grammar; there is no transmission of information from UG to the developing grammar; thus, there is no potential time period in which properties of UG might exist in the mind but have not yet been set within the individual learner's grammar. Thus, if we want to maintain (as indeed I do) that the perceptual system is ultimately delimited by properties of the grammar, then within OT, it must be seen as exerting this force from the very start. We have seen the logical and empirical problems this position encounters above.

There is one more alternative that we might consider in order to maintain that adult perceptual abilities are determined by constraint rankings. According to this position, the constraints in UG would initially be unranked with respect to each other. Thus, *F and PARSE F would both exist in the grammar but neither would outrank the other. At this stage (the default), perception of all contrasts would be possible, since the ranking preventing perception (i.e., *F >> PARSE F) will not have been established; furthermore, we would not expect phonemes to be distinguished at this early stage because the ranking designating the contrastive use of a feature (i.e., PARSE F >> *F) will also not have been established yet.

This hypothesis, of initially unranked constraints, appears to be consistent with the infant perception and L1 phoneme acquisition data; however, it runs askew when it comes to accounting for the adult perceptual abilities. Recall what sort of evidence it is that triggers acquisition of a phonemic contrast: detection of that contrast in the input. Thus, the introduction of a contrast entails ordering PARSE F over *F when a contrast is detected; this is the case of the Chinese speakers. However, if a contrast is not detected, then nothing will occur and the constraints should remain unranked with respect to one another; this would be the case of the Japanese speakers. There would be no evidence for them to set the ranking of constraints to *F over PARSE F, the constraint ranking we saw in the tableaux that the Japanese phonology requires. Moreover, if the constraints remained unranked with respect to one another, this would predict that the Japanese adults (like the Japanese infants) should be able to perceive the /l–r/ contrast. In addition, this position would predict that those unranked constraints would eventually be set given the appropriate L2 input (i.e., presence of a contrast). Therefore, this alternative, though an adequate account of the developmental facts is not viable explanation for the cross-language differences in adult speech perception.

A.3 Conclusions

This short examination OT and discussion of the implications of this theory for first and second language phoneme acquisition have shown that the acquisition and perception findings presented in this dissertation cannot be readily accounted for within the OT framework.

We examined two consequences of the OT premises, in particular: the de-emphasis of segment-internal structure and the move away from viewing underspecification as a property of underlying forms. Both of these notions were central to the theories of phoneme acquisition and L1 phonological interference developed in this dissertation. It is quite clear that the knowledge that children acquire about phonemes is hierarchical in nature and that the hierarchical nature of feature organization (as well as the underspecification of certain features) is reflected in the speech perception system of mature speakers. Thus, any account of these two phenomena must explain why children acquire segments in a uniform order that respects certain relations between features, and why the features that do not contribute to distinguishing contrasts in a given language also do not impinge on the speech perception system of those speakers.

It can be concluded from the discussion that in order to capture the uniformity across children's acquisition of phonemes, and the specific order in which phonemes are acquired, in OT, harmonic rankings between various features must be introduced to the framework. However, if inherent relations between features must be captured with invariant rankings, then it is not entirely clear that this is empirically superior (or at all different, in fact) from encoding dependencies between features in Feature Geometry theory.

Finally, the current means of deriving the underspecificaton of certain features, through the ranking of violable constraints, cannot adequately account for cross-language differences in the perception of non-native segmental contrasts. The constraint hypothesis we investigated above runs into serious logical problems when one considers how this influence of the grammar on speech perception might arise. Thus, when one brings acquisition and speech perception data to bear on the issues of underlying feature organization and underspecification, we are compelled to conclude that the recent move in OT toward de-emphasizing the internal structure of segments and the removal of underspecification of features from underlying representations is premature. The first and second language acquisition data presented in this dissertation receive a more satisfactory explanation under the assumptions of Feature Geometry.