

## INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**UMI<sup>®</sup>**

Bell & Howell Information and Learning  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
800-521-0600



A

INFANT BABBLING AND AMBIENT LANGUAGE

by

DAISY EDMONDSON

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1999

**UMI Number: 9946158**

---

**UMI Microform 9946158  
Copyright 1999, by UMI Company. All rights reserved.**

**This microform edition is protected against unauthorized  
copying under Title 17, United States Code.**

---

**UMI**  
300 North Zeeb Road  
Ann Arbor, MI 48103

© 1999


DAISY EDMONDSON

All Rights Reserved

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

9-23-99  
Date

9-20-99  
Date

  
Chair of Examining Committee

  
Executive Officer

Professor Patricia Brooks

Dr. Darlynn Devenny

Professor Katherine Nelson

Professor Richard Schwartz

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

## Abstract

## INFANT BABBLING AND AMBIENT LANGUAGE

by

Daisy Edmondson

Advisor: Professor Gerald Turkewitz

This dissertation was intended to explore the question of whether infants begin to incorporate aspects of the adult target language to which they are exposed into their own vocalizations prior to the emergence of first words and, if so, whether adult listeners are sensitive to such changes. A further question was which specific aspects of the ambient language emerge earliest in babbling, the phonetic or the prosodic.

In each of eight experiments, English-speaking adult judges were asked to listen to tape-recorded samples of vocalizations of infants and to pick out which infants they believed were from English-speaking homes. Samples were used from 8-month-old Arabic, Chinese, French, and American English-environment infants.

Samples were selected on the basis of duration of utterance alone; there were

no other restrictions on the characteristics of the samples. As a result, samples differed on a number of attributes, including consonant number. In the first study, which contrasted English and Arabic, easy discrimination was found. However, the English and Arabic samples differed widely with respect to number of consonants. Subsequent studies were designed to examine the possibility that consonant number served as the basis for discrimination. Limited support was found for this explanation: Some samples that did not differ significantly with regard to consonant number were discriminated, while others that did differ significantly were not. The possible role of segmental and nonsegmental information in discriminating between the vocalizations of infants from different language backgrounds was assessed by low-pass filtering or scrambling some samples. The filtered samples were discriminated as easily as the full-spectrum versions, while scrambled samples paired with their unscrambled counterparts were not discriminated. The findings that the discriminability of some samples was affected by the context (i.e., that a sample might be more easily discriminated when contrasted with French than with Arabic) and that the masking of specific segmental information through low-pass filtering affected judges' ability to correctly label the samples seem to require an explanation that takes into account the influence of the ambient language.

## ACKNOWLEDGMENTS

I would like, first of all, to acknowledge the contributions of the members of my dissertation committee: Professor Gerald Turkewitz (chair), Professor Katherine Nelson, and Dr. Darlyne Devenny. Professor Turkewitz never gave up on me or my dissertation, even during those times when I came close to giving up myself. He always provided encouragement, advice, feedback, and even criticism--but only when asked. And he never pressured me or gave me unrealistic deadlines. For all this and more, I am truly grateful. Dr. Devenny and Professor Nelson both gave generously of their time and provided me with immeasurable help, both practical and otherwise. The input I received from my outside readers, Professor Richard Schwartz and Professor Patricia Brooks, was invaluable as well.

Many others were helpful to me in the course of completing this dissertation. Professor Marilyn Vihman of the University of Wales at Bangor read and commented upon several earlier drafts and gave me untold help and support through several years of e-mail correspondence. Dr. Bénédicte de Boysson-Bardies provided me with her original tape recordings of French, Arabic, and Chinese infants, without which this research could never have been undertaken. Professor Richard Schwartz, in addition to serving as an outside reader, provided me with tape recordings of American infants. Technical help was provided by Gary Chant of the Ph.D. Program in Speech and Hearing Sciences and Eddy Yeung of the Center for Research in Speech and Hearing Sciences, both at the Graduate Center, as well as by the Educational Resources staff at

the Graduate Center, including Peter Harris (director), John Chianese, and John Wilson. Marie Granthe, Erin Healy, and Toni Wisti, all students at Hunter College, helped with data collection, and Professor Alan Gross and Professor David Rindskopf of the Ph.D. Program in Educational Psychology at the Graduate Center answered countless questions about the statistical analyses. I cannot overlook the members of the dissertation support group I participated in during the last year of dissertation-writing, and especially the facilitator of that group, Dr. Lisa Samstag. Of course, none of this would have been possible without the cooperation of the babies whose vocalizations were used, their parents, and the hundreds of participants who listened to my tapes and gave the task their best effort.

One person who must be singled out for thanks is Barbara Fish, Associate Director of Academic Planning at the Graduate Center, who sat with me hour upon hour, helping me to recreate my lost graphs and feeding me miniature Milky Ways. Her kindness and computer expertise got me through what otherwise might have been a major catastrophe.

Finally, I want to give my thanks to my family--my husband, Stewart Alter, and my children, Imogen and Barnaby--who put up with me through all the ups and downs of completing this project. They provided the distraction--sometimes welcome, sometimes not--that gave a necessary balance to my life during this long and arduous process.

## CONTENTS

ABSTRACT .....	iv
ACKNOWLEDGMENTS .....	vi
LIST OF TABLES .....	ix
LIST OF FIGURES .....	xi
 Chapter	
1. Introduction .....	1
2. Method & Results: Experiment 1 .....	21
3. Method & Results: Experiment 2 .....	34
4. Method & Results: Experiment 3 .....	43
5. Method & Results: Experiments 4 & 5 .....	52
6. Method & Results: Experiment 6 .....	63
7. Method & Results: Experiment 7 .....	78
8. Method & Results: Experiment 8 .....	85
9. General Discussion .....	93
 Appendices	
A. Presentation Orders, Instructions to Participants, Consent Form, Response Sheets .....	110
B. Differences in the Number of Consonants in Sample Sets .....	116
References .....	120

## TABLES

Table	Page
1. Stages of Infant Vocal Development .....	10
2. Observed and Expected Frequencies of Scores (Number Correct) on the English-Arabic Discrimination Task .....	27
3. Number of Consonants in Each English and Arabic Babbling Sample--Experiment .....	32
4. Observed and Expected Frequencies of Scores (Number Correct) on the English-Chinese Discrimination Task .....	36
5. Number of Consonants in Each English and Chinese Babbling Sample--Experiment 2 .....	41
6. Observed and Expected Frequencies of Scores (Number Correct) on the English 1-English 2 Discrimination Task .....	46
7. Number of Consonants in Each English 1 and English 2 Babbling Sample--Experiment 3 .....	50
8. Observed and Expected Frequencies of Scores (Number Correct) on the English 2-Arabic Discrimination Task .....	54
9. Observed and Expected Frequencies of Scores (Number Correct) on the English 2-Chinese Discrimination Task .....	55
10. Number of Consonants in Each English 2 and Arabic Babbling Sample--Experiment 4 .....	61
11. Number of Consonants in Each English 2 and Chinese Babbling Sample--Experiment 5 .....	62
12. Observed and Expected Frequencies of Scores (Number Correct) on the English-French Discrimination Task: Total Group .....	66
13. Observed and Expected Frequencies of Scores (Number Correct) on the English-French Discrimination Task: English-Only Group .....	67

Table	Page
14. Observed and Expected Frequencies of Scores (Number Correct) on the English-French Discrimination Task: French-Exposed Group . . . . .	68
15. Number of Consonants in Each English and French Babbling Sample-- Experiment 6 . . . . .	75
16. Proportion of Correct Identifications of Sample E-A1 Across Experiments 1, 2, 3, and 6 . . . . .	76
17. Observed and Expected Frequencies of Scores (Number Correct) on the Low-Pass-Filtered Babbling Discrimination Task . . . . .	80
18. Observed and Expected Frequencies of Scores (Number Correct) on the Scrambled Babbling-Natural Babbling Discrimination Task . . . . .	87
19. Summary of Results: Experiments 1-8 . . . . .	98
20. Consonant Count per Speech Sample, by Language Group . . . . .	117
21. Log-Transformed Consonant Counts, Means, and Standard Deviations . . .	118
22. Differences Between Sets of Language Samples Tested in Experiments 1-8 . . . . .	119

## FIGURES

Figure	Page
1. Ability of Judges to Discriminate Between English and Arabic Babbling Samples .....	28
2. Number of Times the Origin of Each Sample in Experiment 1 Was Correctly Identified .....	29
3. Mean Number Correct by Order of Presentation in Experiment 1 .....	30
4. Ability of Judges to Discriminate Between English and Chinese Babbling Samples .....	37
5. Number of Times the Origin of Each Sample in Experiment 2 Was Correctly Identified .....	39
6. Mean Number Correct by Order of Presentation in Experiment 2 .....	40
7. Ability of Judges to Discriminate Between English 1 and English 2 Babbling Samples .....	47
8. Number of Times the Origin of Each Sample in Experiment 3 Was “Correctly” Identified .....	48
9. Mean Number Correct by Order of Presentation in Experiment 3 .....	49
10. Ability of Judges to Discriminate Between English 2 and Arabic Babbling Samples .....	56
11. Ability of Judges to Discriminate Between English 2 and Chinese Babbling Samples .....	57
12. Number of Times the Origin of Each Sample in Experiment 4 Was Correctly Identified .....	58
13. Number of Times the Origin of Each Sample in Experiment 5 Was Correctly Identified .....	59
14. Ability of Judges to Discriminate Between English and French Babbling Samples .....	69

Figure	Page
15. Ability of Judges to Discriminate Between English and French Babbling Samples: English-Only Group .....	70
16. Ability of Judges to Discriminate Between English and French Babbling Samples: French-Exposed Group .....	71
17. Mean Number Correct by Order of Presentation in Experiment 6 .....	72
18. Number of Times the Origin of Each Sample in Experiment 6 Was Correctly Identified .....	73
19. Ability of Judges to Discriminate Between Filtered English and Filtered Chinese Babbling Samples .....	81
20. Number of Times the Origin of Each Sample in Experiment 7 Was Correctly Identified .....	82
21. Mean Number Correct by Order of Presentation in Experiment 7 .....	83
22. Ability of Judges to Discriminate Between English and Scrambled English Babbling Samples .....	89
23. Mean Number Correct by Order of Presentation in Experiment 8 .....	90
24. Number of Times the Origin of Each Sample in Experiment 8 Was “Correctly” Identified .....	91

## CHAPTER 1

### Introduction

The studies that will be described here were intended to explore the question of whether infants begin to incorporate aspects of the adult target language to which they are exposed into their own vocalizations prior to the emergence of first words and, if so, whether adult listeners are sensitive to such changes. A further question was which specific aspects of the ambient language emerge earliest in babbling, the phonetic or the prosodic.

#### Infants' Sensitivity to Their Environment

That infants are responsive to various kinds of sensory input from their environment is certainly not news. Indeed, William James's (1890) characterization of the newborn's world as a "blooming, buzzing confusion" has made him a kind of whipping boy for researchers working in the area of infant perceptual development. From birth, infants respond to different tastes in much the same way as adults do (Steiner, 1979; Crook, 1978). Their sense of smell, too, is well enough developed that, within the first week of life, they can discriminate between their own mother's smell and that of another mother (MacFarlane, 1975). Touch may be the earliest-developing sensory system: Even a fetus in the third month of gestation is able to respond reflexively to some kinds of stimulation (Hooker, 1969). Vision, on the other hand, is probably the least developed of the senses at birth. Even so, infants do respond to

visual stimuli, and the visual system develops quite rapidly over the first half-year or so of life (see Aslin, 1987, for a review).

It is the developing auditory system that is of most interest here, however.

Unlike vision--which really cannot begin to operate fully till after birth--audition can and does begin to function prenatally, with consistent responses to acoustic stimuli observed in fetuses starting at 28 weeks (for a review, see Turkewitz, 1988, and Querleu, Renard, Versyp, Paris-Delrue, & Crèpin, 1988). Of greatest importance to this set of studies is the work that has been done in the area of infants' responsiveness to speech stimuli and especially the evidence with regard to the influence of the ambient language

#### The Influence of Language Environment: Perception

There is now abundant evidence that the ambient language influences infants' perceptual abilities very early on--perhaps even prenatally. Even within the first days of life, newborns can recognize familiar voices (DeCasper & Fifer, 1980), stories (DeCasper & Spence, 1986), and melodies (Cooper & Aslin, 1989). This ability to discriminate between the known and the unknown extends as well to familiar and unfamiliar languages. Mehler et al. (1988) tested 4-day-old French and 2-month-old American infants and found that both groups showed evidence of distinguishing utterances in their native language from those in another language, while neither group appeared able to distinguish between utterances from two languages that were both foreign to them. These results were replicated with 2-day-old infants from Spanish- and

English-speaking environments (Moon, Cooper, & Fifer, 1993). Further, in the Mehler et al. study, infants were found to maintain their preference for the native language even when the samples were low-pass filtered, which removes most of the specific phonetic information while preserving the prosodic aspects of the speech. Since other research has shown that it is precisely this information--the prosodic and not the phonetic--that is primarily transmitted to the fetus in utero (Querleu et al., 1988), it appears probable that newborns' preference for their native language, based as it appears to be on prosody, has its roots in prenatal experience. Moreover, the fact that the exposure to the stories and melodies in the studies previously cited (Cooper & Aslin, 1989; DeCasper & Spence, 1986) was solely *prenatal* and the testing for preference *postnatal*, supports the hypothesis that experience begins to influence perception even before birth.

Later in the first year of life, infants develop a greater sensitivity to various language-specific aspects of the speech to which they are exposed. One notable example of a shift from broad-based to language-specific phonetic processing has been identified by Werker (Werker & Tees, 1984; Werker & Lalonde, 1988), who found a decline in infants' ability to discriminate between native and non-native phonetic contrasts occurring around 9-10 months of age.

Other developments that denote an increasing awareness of the target language become evident within the same time frame. Work by Jusczyk and his colleagues has established that 9-month-olds but not 6-month-olds are sensitive to the predominant

English strong/weak stress pattern in speech (Jusczyk, Cutler, & Redanz, 1993), the constraints on ordering of phonetic elements within words (Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993), and the frequency with which certain phonetic patterns appear in English words (Jusczyk, Luce, & Charles-Luce, 1994). Additionally, there is evidence that 9- but not 6-month-olds prefer to listen to speech samples in which pauses have been inserted at points that correspond to major phrasal units (Jusczyk et al., 1992) and that 9-month-olds prefer as well when pauses are inserted at native rather than non-native phrasal boundaries (Jusczyk, 1989). Sensitivity to native as opposed to non-native clausal boundaries, on the other hand, seems to emerge earlier--around 6 months (Hirsh-Pasek et al., 1987; Jusczyk, Kemler Nelson, Hirsh-Pasek, & Schomberg, 1994)--while a similar sensitivity to word boundaries is not apparent till about 11 months (Kemler Nelson, 1989). It is interesting to note that in the Hirsh-Pasek et al. study, the sentences were spoken in motherese. In a follow-up study, Kemler Nelson, Hirsh-Pasek, Jusczyk, and Wright Cassidy (1989) confirmed that the preference for pauses at native clause boundaries was present only when the sentences were presented in infant-directed speech, not in adult-directed speech.

Taken together, these studies support the idea that speech perception is reorganized in ways that are highly predictable from the ambient language. Further, 6 to 9 months of age stands out in these studies as a significant turning point in the development of speech perception.

### Vocal Development in Infancy

Before going on to explore the influence of ambient language on speech production in infancy, it is necessary first briefly to consider what is currently known of normal vocal production in the first year of life.

One early view of babbling, of which Jakobson (1941/1968) was a major champion, was that babbling was not, in fact, a precursor of language but rather a completely separate development, both structurally and functionally. During the babbling period, babies were thought to produce all the sounds of the world's languages. This stage was sometimes followed, according to this view, by a silent period during which babbling ceased; following the silent period, first words would emerge, which marked the beginning of true language.

There is now abundant evidence that this view of the relationship between babbling and language is mistaken. Locke (1983), reviewing data on babbling collected by different researchers in a variety of language groups, noted that, rather than representing an "astonishing . . . diversity of sound productions," as Jakobson (p. 21) claimed, a list of just 11 consonants accounted for 90% of all consonants produced by babies exposed to American English; similar consonant repertoires were reported for infants in non-English language environments. Moreover, study after study has now demonstrated that there is a clear continuity between babbling and word production. The notion of a silent period has not been borne out by research; rather, infants continue to babble even after they have begun to produce words (Vihman, Macken,

Miller, Simmons, & Miller, 1985; Vihman & Miller, 1988). There is also continuity in the sounds used in babbling and in words. Locke (1989) has shown that the same consonants used most frequently in babbling account for 90% of the sounds used in the first 50 words. Other studies have also confirmed a close relationship between sounds in the babbling repertoire and in early words for individual infants (Vihman et al., 1985; Vihman & Miller, 1988).

One reason why it would not be reasonable to expect endless variety in the sounds infants use, either in babbling or in early words, is that infants are operating under certain constraints that limit the sounds they are capable of producing. Kent (1992) points out, as others have before him (e.g., Lieberman, Crelin, & Klatt, 1972), that an infant's vocal tract is more similar to that of a nonhuman primate than to an adult human's. The infant's vocal tract, for example, is smaller overall than an adult's; it is also broader, the pharynx is set higher, and it lacks teeth. The tongue is larger relative to the size of the mouth, and motor control of the tongue is immature. Remodeling of the vocal tract, however, begins early in the first year of life, particularly around 2 to 4 months of age, such that by 6 months, it begins more closely to approximate the adult's vocal tract (Sasaki, Levine, Laitman, & Crelin, 1977). Some of the factors that have been implicated in this remodeling include growth-related changes in the relative size and position of various structures within the vocal tract, as well as the eruption of dentition and molding of the palate in response to mechanical pressure exerted by the tongue during crying, sucking, and swallowing (Thelen, 1991).

Do these anatomical changes, in turn, produce changes in vocal production?

Although, as Kent (1992) maintains, it is not at all clear that relaxation of organic constraints is *sufficient* to produce changes in babbling, it is certainly a *necessary* factor. In dynamic-systems terms, growth and change of the vocal apparatus may destabilize the system and bring about a phase shift, wherein a new attractor state draws the subcomponents into a different stable configuration which permits the emergence of novel behaviors (Thelen, 1991). In the case of infant speech around the middle of the first year of life, the novel behaviors would be new categories of sound, with perhaps the most significant being the onset of what has been called variously canonical babbling (Oller, 1980), reduplicated babbling (Stark, 1980), reduplicated articulatory movements (Koopmans-van Beinum & van der Stelt, 1986), and reduplicated consonant babbling (Roug, Landberg, & Lundberg, 1989).

Detailed descriptions of vocal production over the first year have revealed that development in this domain is stagelike in nature. Whether the ambient language is English (Oller, 1980; Stark, 1980), Swedish (Roug et al., 1989), or Dutch (Koopmans-van Beinum & van der Stelt, 1986), infants apparently proceed through roughly the same stages at roughly the same ages. However, it is important to keep in mind that while the description of stages in vocal development that follows may hold true for the languages listed above, these are all Northern European, and therefore phonologically related, languages, as K. Nelson (personal communication, October 25, 1994) points out. The course of vocal development in other types of languages has not

been carefully charted. Briefly, the stages that have been identified for these Northern European languages--using Oller's (1980) terminology--are as follows:.

*Phonation Stage (0-1 month)*: This period is dominated by what Oller calls quasi-resonant nuclei (QRN), which involve normal phonation (as opposed to breathy voice, for example) but limited resonance.

*GOO Stage (2-3 months)*: Glottal stops appear in this stage, in combination with QRNs, producing a sort of primitive syllable type. These syllables lack the fully resonant nuclei (FRN) and regular timing of mature syllables, however.

*Expansion Stage (4-6 months)*: Marginal babbling, combinations of FRNs and consonant-like elements, occurs in this stage. These syllables, however, still lack the regular syllable timing characteristics of later canonical babbling and tend not to be reduplicated. Raspberries, squeals, and growls are also explored by the infant in this period.

*Canonical Stage (7-10 months)*: During this stage, infants begin to produce syllables containing a FRN and a consonant-like element "in a timing relationship that conforms to mature natural language restrictions" (Oller, 1980, p. 98). Usually, these syllables occur in reduplicated strings, in which the same syllable is repeated (for example, *bababa* or *mamama*). The onset of this stage is generally sudden.

*Variegated Babbling Stage (11-12 months)*: During this stage, infants begin to produce strings of syllables within which the consonant and/or vocalic elements are varied (for example, *daba* or *babi*). The existence of a separate variegated stage has

been disputed, however. Oller (1980), Stark (1980), Koopmans-van Beinum and van der Stelt (1986), and Roug et al. (1989) all described such a stage. However, Mitchell and Kent (1990) found phonetically varied multisyllables at or near the onset of reduplicated babbling.

Although authors differ with regard to some of the details, such as specific ages of onset, there is, overall, very good agreement regarding the major events and their sequence. Table 1 summarizes the four models referred to previously and shows how closely they correspond to one another.

The emergence of mature syllables, as the table shows, is a milestone identified by each of these four investigators. This finding of an abrupt onset of reduplicated babbling sometime between 6 and 10 months has been confirmed in other studies as well--in American English babies (Ramsay, 1984), in Swedish babies (Holmgren, Lindblom, Aurelius, Jalling, & Zetterström, 1986), and in Russian babies (Bel'tyukov & Salakhova, 1973, cited in Locke & Pearson, 1992). Perhaps significantly, the appearance of this pattern in speech production coincides with the appearance more generally of what Thelen (1981) calls rhythmic stereotypies--repetitive movement sequences involving the legs, arms, fingers, hands, and other body parts. Ramsay (1984) has also found that a relationship appears to exist between the onset of babbling and of one-handed reaching. These findings suggest that babbling may reflect an underlying general

Table 1

Stages of Infant Vocal Development

<b>Age in Months</b>	<b>Oller 1980</b>	<b>Stark 1980</b>	<b>K-V B. et al. 1986</b>	<b>Roug et al. 1989</b>
1	phonation	reflexive	uninterrupted	
2	goo		phonation	
3	stage	cooing &		glottal
4	expansion stage	laughter	interrupted	stage
5		vowel play	phonation	velar
6			variation	
7	canonical babbling stage		in	vocalic
8			phonatory	stage
9			domain	
10		reduplicated babbling		reduplicated
11	variegated babbling	non- reduplicated babbling	reduplicated articulatory movements	consonant
12				babbling
13				stage
14				variegated babbling

movement toward rhythmicity, as well as changes in left-hemisphere functioning. The onset of babbling, therefore, in addition to signalling a reorganization in the domain of speech production, may also be a part of a more pervasive developmental reorganization.

One of the defining features of canonical babbling is its "speechiness": Once a baby enters the canonical babbling stage, parents begin to view the baby's vocal productions as "talking" (Oller, 1980). Although the effects of the ambient language on perception are clearly evident some time earlier, as the foregoing review has shown, it is only after this "speechiness" is established that it becomes possible to explore the extent to which the infant's vocal productions approximate those of the adults in the child's environment.

#### The Influence of Language Environment: Production

One way in which the influence of the ambient language is presented, as Vihman and Boysson-Bardies (1994) note, is in the loss of certain sounds that are not supported in the adult language. For example, Vihman (1992) reported a frequent use of *[h]* in the babbling of French infants around 10 months of age--comparable to that of the English infants--although the sound is hardly ever used in adult French. By the time these same infants reached the 25-word point, however, the use of the sound had dropped dramatically among the French infants, while the level of use dropped only slightly among the English infants. Another example from Vihman's (1993) work is the case of a French child who produced 41% of her vocalizations with a final consonant, in contrast to the other French children in the study, who produced a mean of only 6%

with a final consonant. Over time, this child's production of final consonants dropped to 1%--a decline that is probably attributable to the fact that French is not a language that contains a high percentage of consonant-final words (only 26%, in comparison with English's 61%).

The other way in which ambient-language effects can be sought, according to Vihman and Boysson-Bardies (1994), is in the presence, rather than absence, of certain features. There are three types of effects that have been studied: effects on prosody, on vowel space, and on place of articulation of consonants. In each case, studies have shown a general movement toward the target language, beginning within the canonical babbling stage.

Given the salience to infants of the distinctive prosody of infant-directed speech, which is demonstrated from birth in a preference for infant-directed over adult-directed speech (Cooper & Aslin, 1989), it is not unreasonable to expect that changes in prosody might be the first manifestation of the influence of the ambient language. Indeed, there are numerous indications in the literature that this is the case. Many of the cross-language studies of prosody that have been done to date have looked at children over one year of age (e.g., Boysson-Bardies, Bacri, Sagart, & Poizat, 1981; Boysson-Bardies, Sagart, & Bacri, 1981; Hallé, Boysson-Bardies, & Vihman, 1991; Vihman, DePaolis, & Davis, 1998), but even in younger children, differences have been detected. Intonational patterns in reduplicated multisyllabic vocalizations have been compared in English- and French-learning infants of 6-12 months and have been found

to differ in proportion of rising vs. falling pitch contours (Whalen, Levitt, & Wang, 1991)--a pattern that corresponds to the major prosodic difference between the two adult languages (Delattre, 1961, cited in Vihman & Boysson-Bardies, 1994).

Differences in final-syllable lengthening, timing of non-final syllables, and number of syllables per utterance have also been found in contrasting the vocalizations of French and English infants between 4 and 17 months (Levitt & Wang, 1991). Again, these differences reflect prosodic differences between the two adult languages as well (see Levitt & Wang, 1991, for a review).

A number of studies have used adult discrimination as a measure of whether and to what degree infants' vocalizations differ as a function of language environment. These studies have shown mixed results. Tervoort, for example, in an unpublished study (cited in Boysson-Bardies, Sagart, & Durand, 1984), reported that Dutch adults could identify the vocalizations of Dutch infants starting around 6 months. Weir (1966), contrasting vocalizations from American and Chinese 5- to 7-month-old infants, also found that discrimination was possible. Atkinson, MacWhinney, and Stoel (1968), on the other hand, were unable to replicate Weir's results, using her original samples. Likewise, Olney and Scholnick (1976) did not find discrimination between babbling samples of American and Chinese infants of 6, 12, and 18 months--although, as Boysson-Bardies et al. point out, the design of this study presented several methodological problems, among them: 1) adult imitations of infant vocalizations were included; 2) the Chinese infants used were growing up in the United States and might,

therefore, have been exposed to English; and 3) shorter vocalizations were spliced together to make up the 15 second samples. Nor was Kuehn (1985) able to find evidence that American or Chinese adults could discriminate between vocalizations of American and Chinese infants prior to a year of age, although they did show the ability to do so starting at one year.

Perhaps the most extensive of the studies to use an adult-discrimination paradigm to look at cross-linguistic differences in infant vocal production was conducted by Bénédictte de Boysson-Bardies and her colleagues (Boysson-Bardies et al., 1984). It is this study that served as the inspiration for the experiments that make up this dissertation, which will be described in the chapters that follow.

In their first experiment, Boysson-Bardies et al. (1984) asked 40 native French-speakers to listen to pairs of babbling samples that had been collected from 8- and 10-month-old Arabic, Cantonese Chinese, and French infants. These three languages were chosen because they are "strongly contrasted along the dimensions of laryngeal and supralaryngeal settings ... and that of prosody" (p. 4). Thirty-two samples, each 15 seconds long, were selected. The criteria used for selection of sequences were: 1) duration of at least 15 seconds; and 2) pauses no longer than 1,500 milliseconds. Five of the samples were composed of two shorter sequences spliced together, separated by a pause of less than 1,500 milliseconds. Altogether, eight samples were selected from four 8-month-old French infants and eight from four 10-month-old French infants. For the Arabic and Chinese infants, four samples were selected from

two children from each language group at each age. The adults were told that one sample in each pair came from a French infant and the other from a non-French infant. They were asked to identify which one was the French infant. The results showed that these adult judges were very good at identifying the French infants at both ages when French was paired with Arabic (a mean of 75.8% correct identifications for the 8-month-olds and 74.4% at 10 months). For the French-Cantonese pairings, however, the results were less clear: For the 8-month samples, the judges did rather well (69.4% correct); for the 10-month samples, performance was much worse (31.9%).

In a second experiment, three groups of native French-speaking adults were employed as judges. Those in Group 1 were not told the origin of the non-French samples; those in Group 2 were told the origin of the non-French babies; and those in Group 3, who were trained phoneticians, were also told the origin of the non-French samples. In this experiment, samples from three ages--6, 8, and 10 months--were used, but only the French-Arabic pairing was tested, as new 6-month-old Cantonese samples could not be obtained at the time. The same sample-selection criteria as in Experiment 1 were used for selection of the samples from 6- and 8-month-olds; for the 10-month-olds in Experiment 2, however, only sequences containing highly articulated productions of reduplicated babbling were used. The only group of judges who exceeded chance in distinguishing among the 6-month-old infants were the phoneticians (68% correct), while all three groups performed at significantly better than chance levels with the 8-month-olds (74% for Group 1; 67.5% for Group 2; 72.2% for Group 3). The surprise

came in that no group was able to classify the samples at above-chance levels at 10 months (53.3%, 49%, and 56.9% correct for the three groups, respectively). The authors concluded that this inability to classify the 10-month samples was due to the fact that these consisted primarily of "sequences of isolated or reduplicated syllables [that were] generally poor in intonational patterns" (p. 9). The success at 8 months, therefore, was attributed to the nonsegmental cues in the infants' babbling.

From these results, Boysson-Bardies inferred that a general tendency exists for babbling to incorporate some of the distinctive prosodic features of the adult language. This led her next to turn her attention to vowel formants, as it is vowels rather than consonants that carry much of the suprasegmental information in speech (Polka & Werker, in press). Once again, she and her colleagues (Boysson-Bardies, Hallé, Sagart, & Durand, 1989) looked at infants from French, Chinese, and Arabic language environments, but this time they added an English group as well. All infants were recorded once, at 10 months of age. Previously, researchers had been unable to find any statistically significant differences between the vowels used by infants from different language groups when phonetic transcriptions were used (e.g., Oller & Eilers, 1982). However, the acoustic analyses used in this study uncovered significant differences among the groups, with the English babies producing relatively more front vowels, the French infants more mid central vowels, and the Chinese and Arabic babies more low central ones.

More recently, Boysson-Bardies and Vihman (1991) have explored differences

in consonantal place and manner categories for French, English, Japanese, and Swedish infants, between the ages of approximately 10 and 17 months. Again, as mentioned above, previous analysis had found that the babbling repertoires of infants from different linguistic environments were virtually identical (Locke, 1983). In this study, however, the distribution of consonant *tokens*, rather than *types*, was the measurement of interest. Looked at this way, differences emerged in both categories, and these differences mirrored those found in the target languages. For example, even as early as 10 months, the English and French infants used many more labials in their babbling than did the Japanese and Swedish infants; the French and Japanese infants used more nasals than the infants from other language groups; and the babbling of the Swedish and French infants contained more liquids than did the babbling of English and Japanese infants. (Not all of these differences were statistically significant, however.)

In summary, then, the notion of "babbling drift" proposed by Brown more than 40 years ago (Brown, 1958) has received considerable support from subsequent research. Moreover, there is independent evidence that perception, too, "drifts" in the direction of the target language, as shown in the earlier review. Intuitively, it seems that a relationship must exist between perception and production in this period prior to the emergence of first words, but it is not clear, in detail, what that relationship might be. We know from work with deaf infants that they enter the canonical babbling stage later than do hearing infants (Lynch, Oller, & Steffens, 1989; Oller & Eilers, 1988; Oller, Eilers, Bull, & Carney, 1985; see also Kent, Osberger, Netsell, & Hustedde, 1987).

We also know that there are distinct differences in the consonantal repertoires of hearing and non-hearing infants, with the non-hearing infants having smaller repertoires that decrease over time and fewer multisyllabic utterances containing true consonants (Stoel-Gammon & Otomo, 1986). Hearing-impaired infants also apparently show very high proportions of labials in their repertoires after 15 months of age, possibly because of visual cuing (Smith, 1982). Such differences indicate that perception is essential to vocal development. The findings of the studies of Boysson-Bardies and her colleagues, described above, also suggest a leading role for perception, in that infants must perceive the sounds of the adult language in order to attempt to match them. Just what might be the reciprocal influence, if any, of production on perception is something about which there is speculation but little evidence one way or the other. One missed opportunity to look directly at a perception-production relationship is embodied in a study by Locke and Pearson (1990) of an infant who was tracheostomized between 5 and 20 months. Following decannulation, her vocal development was found to be markedly delayed, which indicates that maturation alone cannot account for babbling onset. It would have been interesting to know whether there were any effects in the perceptual realm, but no testing of this type was done.

### The Present Research

At the outset, it was stated that the studies reported here were done with two goals in mind: First, to attempt to find out how early on infants begin to incorporate

elements of the adult target language into their babbling; and, second, to ascertain whether the earliest-appearing elements would be prosodic or phonetic ones. More specifically, these studies aimed to replicate and extend Boysson-Bardies et al.'s (1984) findings with regard to the ability of adult judges to discriminate between infants of differing language backgrounds and to explore the possible bases upon which such a discrimination might be made.

The question might be asked, Why use the adult-discrimination paradigm when acoustic analyses of infants' vocalizations have yielded direct evidence of movement toward the ambient language? The answer is that, while direct analyses of babbling may provide information about what infants are actually doing, it is quite possible that adult listeners may pick up on these changes later or perhaps earlier than they can be detected by acoustic analysis. The point at which these changes are detectable by adult listeners is important in that--as with the emergence of canonical syllables, which, as described above, leads parents to view their infants' vocalizations as "talking" (Oller, 1980)--the discernible "drift" of an infant's babbling toward the target language is likely to lead parents to respond differently as well.

The following eight experiments, which will be described in detail in the chapters that follow, were conducted:

Experiment 1: American English-speaking judges were tested on their ability to discriminate between the vocalizations of American infants being raised in English-speaking households and Arabic-environment infants in an effort to extend Boysson-

Bardies et al.'s (1984) findings.

Experiment 2: English-speaking judges were tested on their ability to discriminate between American and Chinese infants, again to extend the earlier findings.

Experiment 3: The original set of English samples was contrasted with a new set from same-age infants to see if number of consonants in the samples might influence discrimination.

Experiment 4: The new set of English samples was contrasted with Arabic, again to see what role consonants might play in discrimination.

Experiment 5: The new set of English samples was paired with Chinese, with the same purpose in mind.

Experiment 6: The original set of English samples was contrasted with French samples that contained fewer consonants to see if a consonants-only explanation would hold.

Experiment 7: The original English and Chinese samples were low-pass filtered to attempt to isolate the contribution of specific segmental information to discriminability.

Experiment 8: The original English samples were scrambled and then paired with their unscrambled counterparts to attempt to isolate the contribution of nonsegmental information to discriminability.

## CHAPTER 2

### Method & Results

#### Experiment 1: English-Arabic Discrimination Task

The purpose of this first study was to extend Boysson-Bardies et al.'s (1984) findings by testing English-speaking judges' ability to discriminate between the babbling of infants from English and Arabic environments. Unlike that earlier study, where 6- and 10-month-old infants were also included, only 8-month-old infants were used in this and the studies that follow, as that was the age at which Boysson-Bardies et al. (1984) got the clearest evidence of discrimination.

#### Participants

Fifty American English-speaking judges (9 male, 41 female) were recruited by word-of-mouth. Although virtually all had studied a second language in school, anyone who identified a language other than English as his/her first language was excluded. None of the participants had studied Arabic, and none had any known hearing impairments. Fifteen of the judges were parents, 13 were not, and 22 did not provide this information; 28 had some phonetic training. Twenty-three of these were students in a phonology class, which accounts for the high number of participants with phonetic training in this study as compared with the others that follow. Although participants in this study were not asked to list their age nor whether or not they had received any musical training--two items that were added to the response sheet in later studies--in the case of 6 of the 50 participants I was able to find out the age and in 9 the

information about musical training at a later date. The mean age of those for whom this information was available was 50.5 years (range = 39-70); 8 of the 9 participants asked did have some musical training.

### Materials

Tape recordings of four normally developing 7½-month-old infants from English-speaking homes were made as part of another study of early phonetic inventories (Petinou, Schwartz, Mody, & Gravel, in press). Each infant's vocalizations were digitally recorded in a sound-treated test booth using a Telex Wireless clip-on microphone. Sessions lasted approximately 20 minutes, during which time infants were encouraged to vocalize during semi-structured play activities. The four infants used in the present study were selected by one of the researchers of the phonetic-inventory study from among the eight normally developing infants participating in that study.

From these recordings, every babbling segment that was approximately 10 seconds in length was identified. Initially, 15 second segments were sought--in line with those used by Boysson-Bardies et al. (1984). However, very few segments met that criterion, so the criterion was changed to broaden the pool from which segments could be chosen. Four samples were then selected randomly from among those that met the 10-second criterion. Because of differences among the infants with regard to vocal productivity, it turned out that the four samples came from just two of the four infants--two samples from each. The sex of the infants was not known during selection.

The Arabic samples were selected from tapes provided by Bénédicte de

Boysson-Bardies. These recordings were of the same infants used in her earlier study (Boysson-Bardies et al., 1984), though the segments selected for use in the current experiment were not identical to those selected for use in the Boysson-Bardies et al. study as the criteria differed somewhat. (See pages 14-15, Chapter 1, for further details on the selection criteria used by Boysson-Bardies et al.) Infants' vocalizations were recorded in their homes with a Sony 510/2 tape recorder and an MD 441N Sennheiser microphone. Sessions lasted approximately 30 minutes. In that study, only two 8-month-old Arabic infants were recorded; therefore, two babbling samples of approximately 10 seconds in length were chosen at random from each of the two infants. Again, the infants' sex was not known during selection.

The eight samples--four English, four Arabic--were then used to prepare three audio tapes. The samples were combined on the tapes in three different random orders (see Appendix A). Silences of approximately 10 seconds were inserted between samples. Two Marantz tape recorders (Model PMD222), connected by an audio cable, were used to create the tapes.

### Procedure

The task the judges were asked to perform was to identify which of the samples on the tape they were listening to derived from American infants and which from non-American infants. Participants were told the approximate age of the infants on the tape. Further, they were told that some of the infants were growing up in households where English was spoken and some were growing up in households where some

language other than English was spoken. The other language was not identified.

Participants were permitted to listen to the tape as many times as they chose, with twice being the minimum. (Those tested in the classroom situation, described below, were only able to listen to the tape twice.) The full text of the instructions given to participants can be found in Appendix A, along with the written consent form each was given to read and sign prior to testing.

Most participants were tested individually in a quiet room, listening to the tape on a portable tape player (Aiwa HS-T19) with mini-headphones (Aiwa A081). Some ( $n = 23$ ), however, were tested in a classroom situation with the samples played directly from a tape recorder. These were the 23 students from the phonology class. Eight judges were tested on order A, 32 on order B, and 10 on order C. The reason for this inequality is that the 23 participants who were tested in the classroom setting (the phonology students) all, of necessity, heard the same tape, order B.

Responses were recorded on a sheet provided for this purpose (see Appendix A). On the same response sheet, information on other background variables was also requested. For "language background," participants were asked to list any languages they had studied, for how long, and at what level (i.e., in junior high, high school, or college). Judges were instructed to answer "yes" to "phonetic training" if they had ever had a course in speech science, phonetics, phonology, or a related area. For "children," judges were asked to answer yes or no. Participants' sex was also recorded. Visual examination of the data did not indicate that any of these background factors—nor any

of the others added in later studies—contributed to the outcome of this and subsequent studies. Further analyses were not performed, as there were too few data points and too many variables for such analyses to be valuable.

Professor Marilyn Vihman, then at Southeastern Louisiana University, collected the data from the 23 phonology students; Toni Wisti, an undergraduate at Hunter College, also provided some assistance with data collection.

### Results

For each participant, a score between 0 and 8 correct was possible. A correct response was defined as correctly identifying a sample as English or non-English. In this study, however, it was the potential discriminability of samples from the two language backgrounds that was of principal interest, more than the ability to correctly identify which samples were English. In terms of pure discrimination, the judge who correctly identifies all eight samples is equal to the judge who is wrong on all eight: Both are identifying two distinct groups; one merely has the labels reversed. The binomial probability of either case is identical. Therefore, a decision was made to collapse the response categories as follows: the number of judges getting 0 or 8 correct were combined, as were the judges getting 1 or 7 correct, 2 or 6 correct, and 3 or 5 correct; those performing at chance constituted a category unto themselves. Because of the very small number of cases that would be expected in some of these categories by chance, the categories were then further collapsed to the following three: the first containing the number of judges receiving a score of 0, 1, 7, or 8; the second consisting of those who

received scores of 2, 3, 5, or 6; and the third, those who got a score of 4. The observed and expected frequencies--from a table of binomial probabilities--of cases falling into each of these three categories is shown in Table 2.

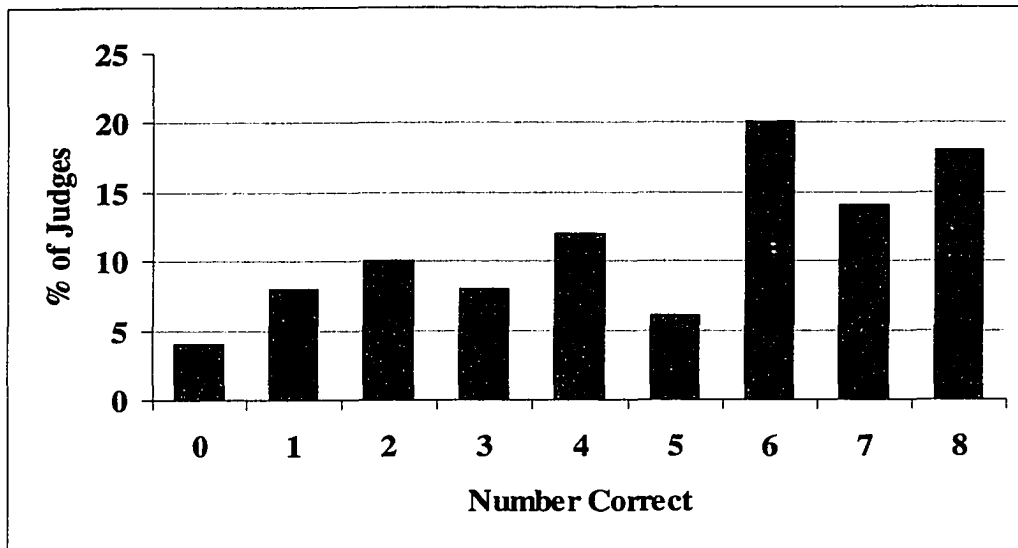
It is clear from looking at the distribution of scores that far more judges scored in the first category than would be expected by chance, while far fewer scored 2, 3, 4, 5, or 6 correct than expected, a result that was statistically significant as well as immediately apparent,  $\chi^2 (2) = 105.27, p < .001$ . Not only were the judges in this study clearly discriminating the two language groups from each other; they got the labels right, as well: As Figure 1 illustrates, more than half of the judges correctly identified the origin (i.e., English or not) of six (75%) or more of the babbling samples.

The possibility that some individual babbling samples might be easier to discriminate than others was checked by performing a chi-square on the frequencies with which samples were correctly identified as English or not by the 50 judges (shown in Figure 2). The results displayed no evidence that some samples were easier to categorize by language group,  $\chi^2 (7) = 4.99$ . There was evidence of an order effect: The mean number correct for order A was 3.87, for order B it was 5.59, and for order C 3.6; a one-way ANOVA performed on these means was significant ( $F [2, 47] = 3.69, p < .05$ ; Figure 3). Additionally, female participants scored significantly higher than did males: the mean number correct for female judges was 5.41, while for males it was 2.66 ( $t [48] = 3.39, p < .001$ , two-tailed). Phonetically-trained judges scored higher than did non-phonetically-trained ones, though the difference was not

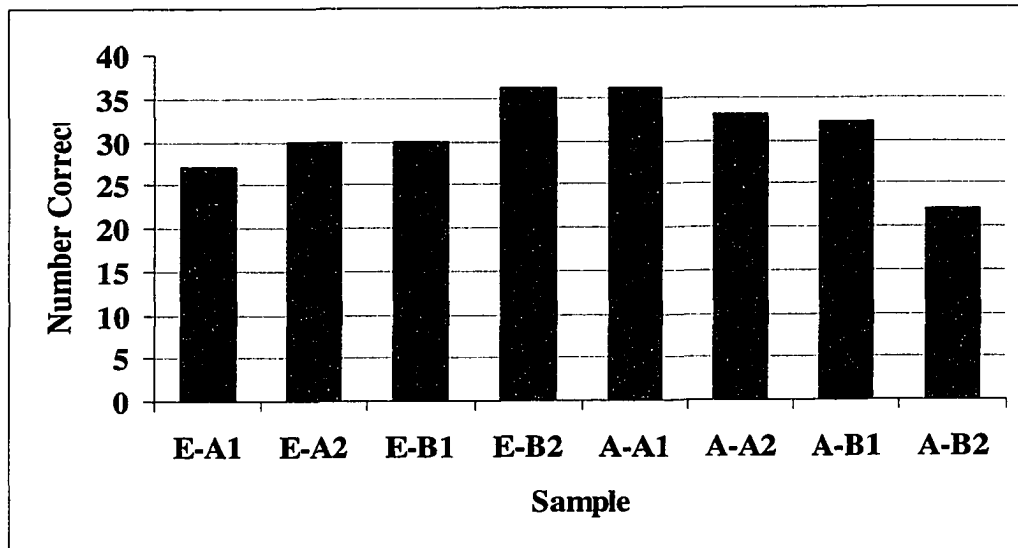
Table 2

Observed and Expected Frequencies of Scores (Number Correct) on the English-Arabic Discrimination Task ( $N = 50$ )

	0, 1, 7, or 8	2, 3, 5, or 6	4
Observed	22	22	6
Expected	3.51	32.82	13.67



**Figure 1.** Ability of judges to discriminate between English and Arabic babbling samples ( $N = 50$ ).



**Figure 2.** Number of times the origin of each sample in Experiment 1 was correctly identified ( $N = 50$ ).

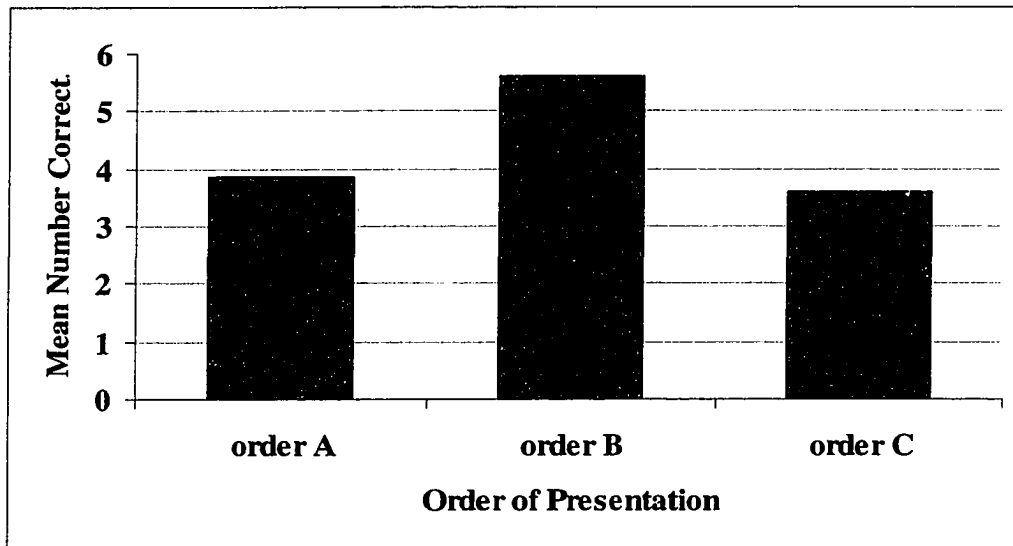


Figure 3. Mean number correct by order of presentation in Experiment 1.

statistically significant ( $M = 5.3$  for the phonetically-trained judges,  $M = 4.5$  for the other judges;  $t [48] = 1.03$ ).

### Discussion

Overall, these results support Boysson-Bardies et al.'s (1984) finding that babbling differs by target language. However, closer scrutiny of the samples used raises a question as to the basis for the discrimination. Boysson-Bardies et al. (1984) concluded that language-specific prosodic differences were what made the infants' babbling discriminable. When the samples used in the present study were examined for evidence of canonical babbling, however, another potentially important difference emerged: While all of the English samples contained abundant consonant sounds, none of the Arabic samples contained a single consonant (Table 3). At this point, therefore, it was unclear whether the judges' discrimination was made on the basis of language-specific properties of the babbling or merely on the basis of the presence or absence of consonants in the samples.

One way to answer this question might have been to contrast the English samples with different Arabic samples--samples containing more consonant sounds. The assumption here is that the difference between the samples was not truly a language-specific one. There is no a priori reason to believe that there would, in general, be fewer consonants in Arabic-environment infants' babbling than in the babbling of English-environment infants. Of course, this is a question that can only be

Table 3

Number of Consonants in Each English and Arabic Babbling Sample--Experiment 1

	E-A1	E-A2	E-B1	E-B2	A-A1	A-A2	A-B1	A-B2
#	6	8	15	12	0	0	0	0

answered through a closer study of the course of phonological development in Arabic-exposed infants, something that is beyond the scope of the current investigation.

In the absence of additional Arabic samples, a new study was undertaken—Experiment 2, which is described in the next chapter. In this study, the same English samples used in Experiment 1 were contrasted with babbling samples from infants of another language group.

Two additional concerns that needed to be addressed in the next study were the order effect and gender difference in performance noted in Experiment 1. It is possible that both of these can be accounted for by the subgroup of 23 students from the phonology class. All of these students—who performed above the mean, as noted previously—were tested on order B, which, as can be seen in Figure 3, had the highest mean performance of the three orders ( $M = 5.59$ , as compared with 3.87 for order A and 3.6 for order C). Moreover, all of these 23 students were female, which boosted the mean of the female judges over that of the males, of which there were very few to begin with. Notwithstanding the likelihood that this subgroup of phonology students was responsible for these unexpected results, a decision was made to test equal numbers of male and female judges in Experiment 2, as well as to attempt to balance the number of people tested on each of the three orders.

## CHAPTER 3

### Method & Results

#### Experiment 2: English-Chinese Discrimination Task

This experiment was designed to clarify whether the order and gender effects noted in Experiment 1 were genuine or merely artifacts of the uneven subject groups and to take another look at the role that consonants might be playing in the discrimination of samples. Additionally, as in Experiment 1, this study sought to further expand upon Boysson-Bardies et al.'s (1984) results, this time by testing English-speaking judges' ability to distinguish between American and Chinese (Cantonese-background) 8-month-olds' babbling.

#### Participants

Thirty-two American English-speaking judges (16 male, 16 female) were recruited by word-of-mouth. None had participated in Experiment 1. Indeed, no judge in any of the eight studies participated in more than one experiment. As in Experiment 1, nearly all had studied a second language, but none had studied Chinese; none had a known hearing impairment; and none identified a language other than English as his/her first language. Seven were parents, and none had any phonetic training. Information on age and musical training was not collected from participants in this study.

#### Materials

The same American babbling samples were used in this experiment as were used in Experiment 1.

The Chinese samples were selected in the same manner as described in Experiment 1. Once more, these samples came from tapes provided by Bénédicte de Boysson-Bardies and were the same infants used in her earlier study (Boysson-Bardies et al., 1984). Again, since only two Cantonese-background 8-month-olds, of unknown sex, were available, two babbling samples of approximately 10 seconds in length were chosen at random from each. These four samples were then re-recorded with the English samples onto tapes as described in Experiment 1, using the same three random orders (see Appendix A).

#### Procedure

The procedure was identical to that described in Experiment 1, with the exception that, in this experiment, *all* participants were tested individually in a quiet room, listening to the tape on a portable tape player with mini-headphones, the make and model of which are not available. Twelve judges were tested on order A, 8 on order B, and 12 on order C.

All of the testing for this study was done by Marie Granthe, an undergraduate at Hunter College.

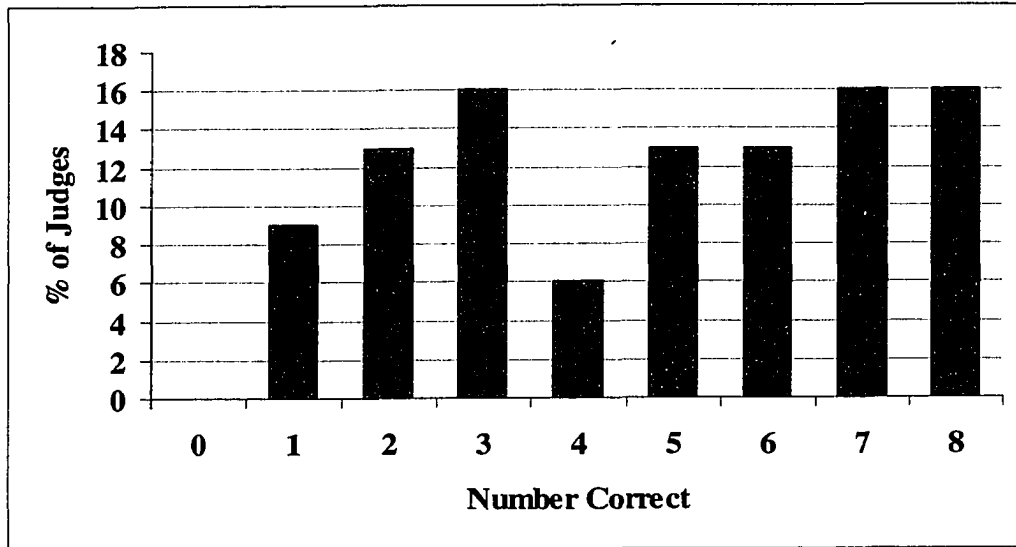
#### Results

The same analyses were performed on the data in this experiment as were employed in Experiment 1. The observed and expected frequencies of scores are displayed in Table 4 and a graph of the distribution in Figure 4. As in the previous experiment, the difference between observed and expected frequencies was statistically

Table 4

Observed and Expected Frequencies of Scores (Number Correct) on the English-Chinese Discrimination Task ( $N = 32$ )

	0, 1, 7, or 8	2, 3, 5, or 6	4
Observed	13	17	2
Expected	2.25	21	8.75



**Figure 4.** Ability of judges to discriminate between English and Chinese babbling samples ( $N = 32$ ).

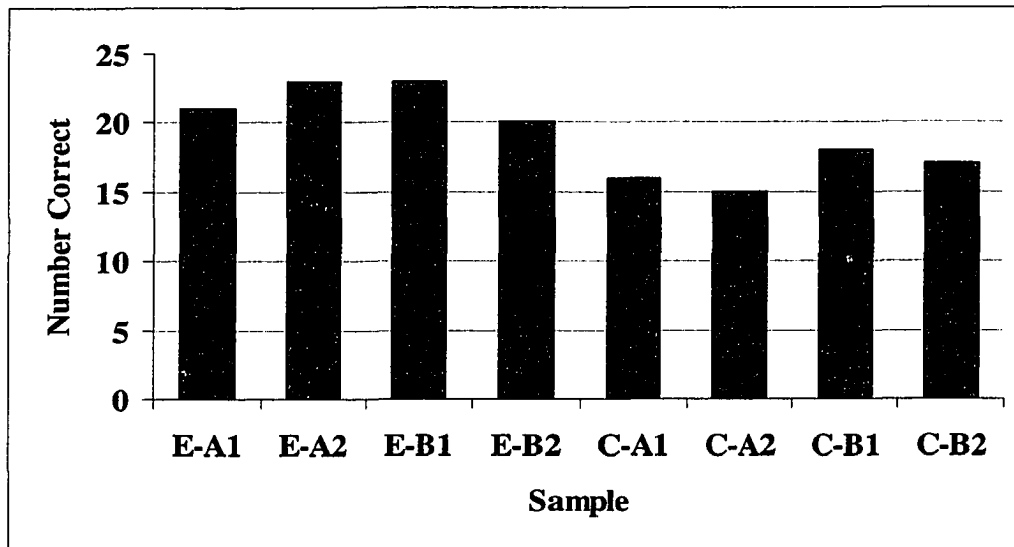
significant ( $\chi^2 [2] = 67.25, p < .001$ ), and nearly half of the judges scored 6 (75%) or better on the task.

With these samples, as with those used in Experiment 1, there was no evidence that some were easier to categorize than others ( $\chi^2 [7] = 3.49$ ; Figure 5). In contrast to the earlier study, however, there was neither evidence of an order effect (for order A  $M = 4.42$ , for order B  $M = 4.13$ , for order C  $M = 5.58$ ;  $F [2,29] = 1.15$ ; Figure 6), nor of a gender difference (for males  $M = 4.38$ , for females  $M = 5.19$ ;  $t [30] = 0.96$ ).

An ANOVA was performed comparing the mean number of consonants across language groups. The results were significant ( $F [4,15] = 12.49, p < .01$ ). Protected  $t$ -tests were then done to discover which language pairs differed significantly from each other. The mean number of consonants was found to be significantly greater in the English than in the Chinese samples ( $t [6] = 4.8, p < .001$ , two-tailed). The full results of this analysis are given in Appendix B.

### Discussion

It appears that, once again, Boysson-Bardies et al.'s (1984) findings are supported. It still is not clear, however, that language-specific prosodic differences are responsible for these results: Even though the Chinese, in contrast to the Arabic, samples did contain consonants, they contained far fewer than did the English samples (Table 5). In fact, this difference was statistically significant, as noted above. This does not rule out, therefore, the possibility that the *quantity* of consonants--not just their presence or absence--in the samples may have been responsible for the discrimination



**Figure 5.** Number of times the origin of each sample in Experiment 2 was correctly identified ( $N = 32$ ).

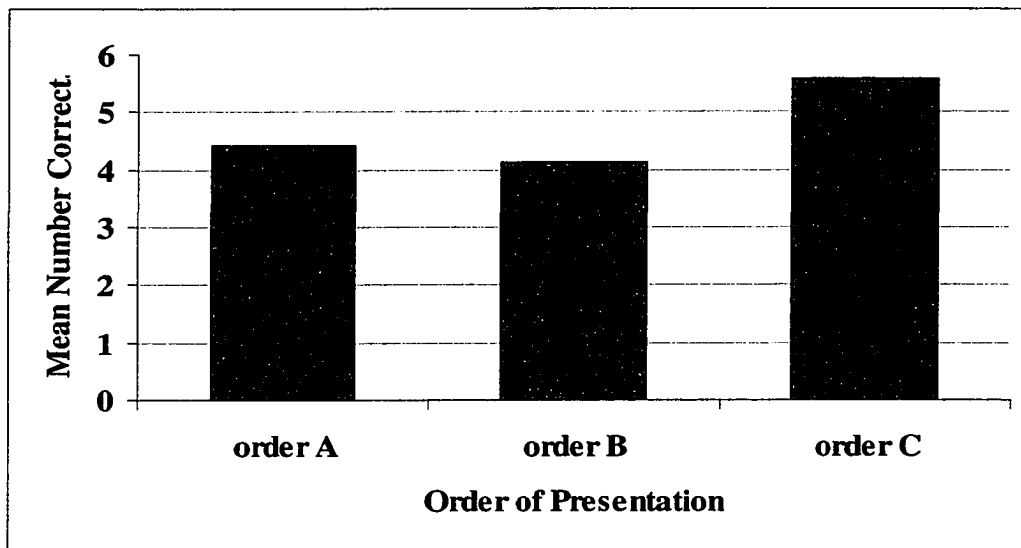


Figure 6. Mean number correct by order of presentation in Experiment 2.

Table 5

Number of Consonants in Each English and Chinese Babbling Sample--Experiment 2

	E-A1	E-A2	E-B1	E-B2	C-A1	C-A2	C-B1	C-B2
#	6	8	15	12	5	2	0	1

found in both Experiment 1 and Experiment 2. The two other questions this study was meant to address, on the other hand--regarding a possible sex difference in performance and an order effect--do appear to have been answered in the negative.

In the study that follows, a further attempt was made to explore the role that number of consonants might play by holding language background constant while manipulating quantity of consonants in the samples.

## CHAPTER 4

### Method & Results

#### Experiment 3: English 1-English 2 Discrimination Task

In this study, the possibility that the quantity of consonants in the babbling samples was responsible for the results of the previous experiments was examined by selecting a new set of babbling samples from 8-month-old American infants and pairing them with the English samples used in Experiments 1 and 2.

#### Participants

Thirty-three American English-speaking judges (4 male, 29 female) were recruited through the participant pool at a local college. The student body of the college from which participants were recruited is approximately 73% female and 27% male. As students signed up for the studies of their choice, there was no practical way to balance for sex in this or any of the studies that follow, all of which made use of the participant pool. The participants' mean age was 19.1 years (range = 17-25). All met the criteria of having no known hearing impairment and of identifying English as their first language. Three had children, 14 had studied music, all were born in the United States, and none had any phonetic training.

#### Materials

The four American babbling samples that were used in the two previous studies were used again in this study. Hereafter, they will be referred to collectively as "English 1."

Two additional 8-month-old American infants from English-speaking households were audiotaped at home in a free-play session with their mothers. A Marantz tape recorder (Model PMD22) with a Shure microphone (Model SM18B) was used to record the infants' vocalizations. Four samples were then selected randomly from among those that met the criterion of being 10 seconds in length, in a manner identical to that used to select the English 1 samples. These new English samples will be called "English 2," to differentiate them from the English samples used previously. The English 1 and English 2 samples were then recorded, as previously described, in the same three random orders that are shown in Appendix A.

#### Procedure

The procedure was identical to that described previously, with all participants tested individually in a quiet room listening to the tape on a portable tape player with mini-headphones (Panasonic RQ-P40 [ $n = 28$ ] or Aiwa TX656 with Aiwa HP-M16 headphones [ $n = 5$ ]). Instructions were unchanged: As in Experiments 1 and 2, participants were told that some of the infants they would be hearing were growing up in non-English-speaking households, though this was not true in this case. Fourteen judges were tested on order A, 12 on order B, and 7 on order C. Inequalities in these groups resulted from exclusion from analysis of data from judges who did not meet the language criterion but who were tested anyway. A revised response sheet was used in this study and in all subsequent studies (see Appendix A). In addition to the information recorded before, starting with this study the following information was recorded as well:

age, musical training, place of birth, headphone position (normal or reversed), equipment used, and number of times participants listened to the tape.

### Results

In this study, identifying English 1 as “English” and English 2 as “non-English” were defined as “correct” responses. The data were analyzed in the same way the data in the previous experiments had been, with the same result: The difference between the observed and expected frequencies of scores (shown in Table 6) was, once again, statistically significant ( $\chi^2 [2] = 10.95, p < .01$ ) and nearly half of the judges scored 6 (75%) or better on the task (Figure 7). No differences were found in the discriminability of the samples ( $\chi^2 [7] = 3.03$ ; Figure 8), nor was there an order effect (order A  $M = 5.21$ , order B  $M = 5.33$ , order C  $M = 5.86$ ;  $F [2, 30] = 0.41$ ; Figure 9). A gender effect, on the other hand, was evident in this experiment (for males  $M = 6.75$ , for females  $M = 5.2$ ;  $t [31] = 2.01, p < .05$ , two-tailed). In view of the fact that there were only four male participants in this study, this result is suspect. There was a significant difference in the mean number of consonants in the two sample sets ( $t [6] = 4.24, p < .001$ , two-tailed; see Appendix B).

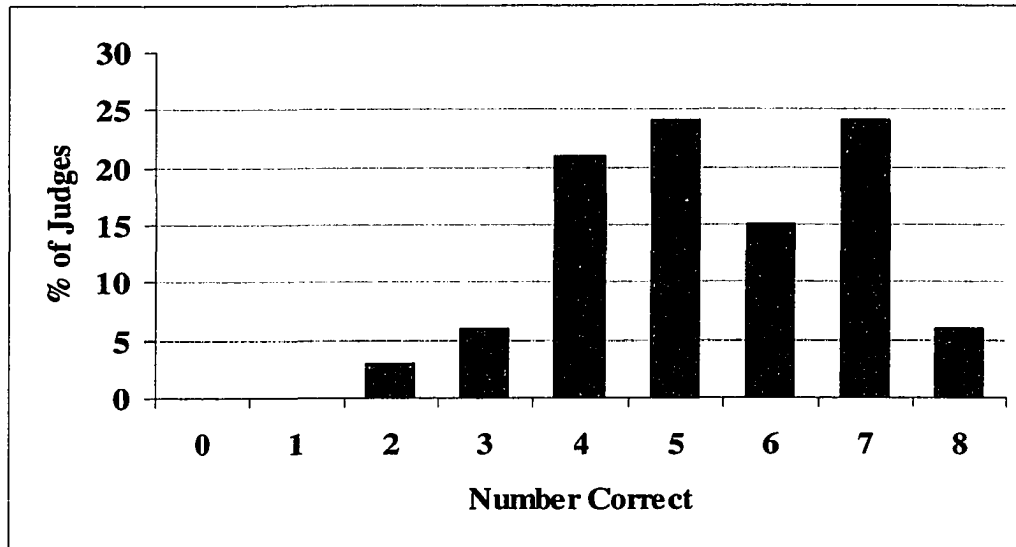
### Discussion

These results show that the English 1 samples and the English 2 samples were discriminated from each other and that, moreover, participants most often judged English 1 to be “English” and English 2 to be “not English.” In Table 7, the number of consonants per sample is shown, and it can be readily seen that English 1 contains far

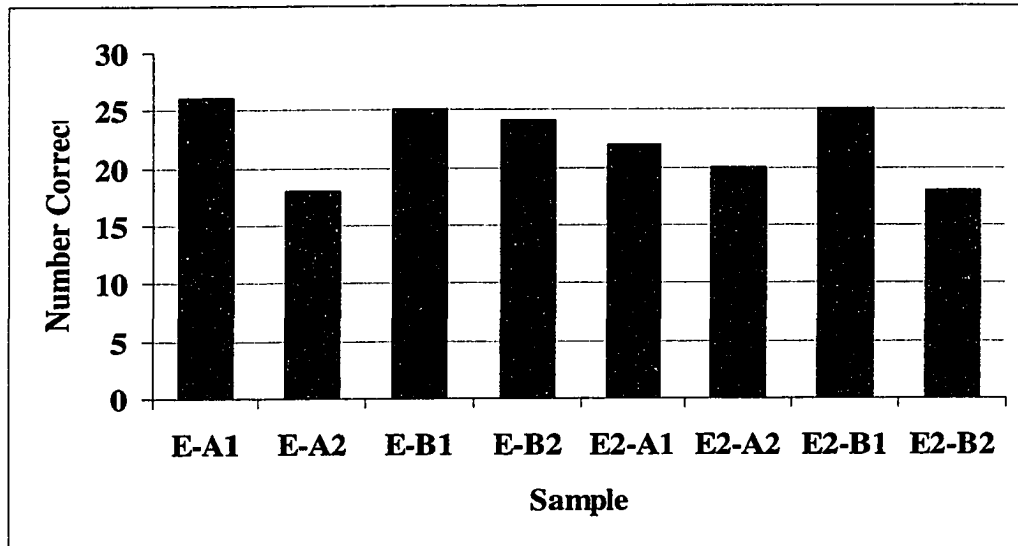
Table 6

Observed and Expected Frequencies of Scores (Number Correct) on the English 1-English 2 Discrimination Task ( $N = 32$ )

	0, 1, 7, or 8	2, 3, 5, or 6	4
Observed	10	16	7
Expected	2.25	21	8.75



**Figure 7.** Ability of judges to discriminate between English 1 and English 2 babbling samples ( $N = 33$ ).



**Figure 8.** Number of times the origin of each sample in Experiment 3 was “correctly”<sup>1</sup> identified ( $N = 33$ ).

---

<sup>1</sup>“Correctly,” in this case, means identifying English 1 as “English” and English 2 as “not English.”

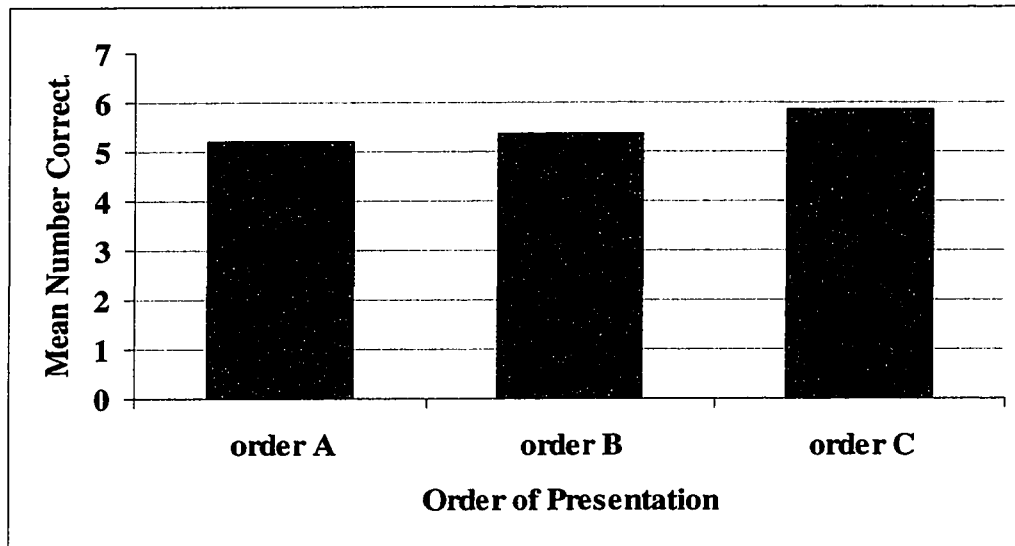


Figure 9. Mean number correct by order of presentation in Experiment 3.

Table 7

Number of Consonants in Each English 1 and English 2 Babbling Sample--Experiment 3

	E-A1	E-A2	E-B1	E-B2	E2-A1	E2-A2	E2-B1	E2-B2
#	6	8	15	12	1	1	0	1

more consonants than does English 2, a difference that is statistically significant.

Since all the samples--both English 1 and English 2--were derived from 8-month-old American infants living in English-speaking households, a marked prosodic difference between the two sets of samples would not be expected. It seems, therefore, that the difference in number of consonants between English 1 and English 2--which appears to be a function of sampling-- is the most likely explanation for the judges' ability to discriminate between them. The next two studies were designed to further explore the possibility that the discrimination found in the previous studies resulted from differences in the number of consonants rather than from language-specific differences in the babbling.

## CHAPTER 5

### Method & Results

#### Experiment 4: English 2-Arabic Discrimination Task &

#### Experiment 5: English 2-Chinese Discrimination Task

In these next two experiments, the English 2 samples were paired with the Arabic samples that were used in Experiment 1 and with the Chinese samples that were used in Experiment 2 to see if the discrimination that was found in those studies could be replicated with a set of English samples containing fewer consonants.

#### Participants

For Experiment 4, 39 participants (6 male, 33 female) were recruited through the participant pool at a local college. Their mean age was 22.9 years (range = 14-49); six judges did not supply this information. All met the criteria of having no known hearing impairment and of identifying English as their first language. Six had children, 19 had studied music (this information is missing for one participant), all were born in the United States, and none had any phonetic training.

For Experiment 5, an additional 39 judges (14 males, 25 females) were recruited from the participant pool. Their mean age was 21.7 years (range = 17-40). All met the hearing and first-language criteria. Five were parents, 20 had studied music, all were U.S.-born, and none had phonetic training.

#### Materials

In Experiment 4, the English 2 samples from Experiment 3 were paired with the

Arabic samples from Experiment 1; in Experiment 5, English 2 was paired with the Chinese samples used in Experiment 2. As no order effect was expected, the stimuli in each of these two experiments were presented in only one order: order B for the English 2-Arabic pairing and order A for the English 2-Chinese pairing (see Appendix A).

### Procedure

The procedure remained unchanged in both of these experiments, with all testing done individually in a quiet room using a portable tape player with mini-headphones (Panasonic RQ-P40).

### Results

A chi-square analysis was again performed using observed and expected frequencies of scores, collapsed into three categories, as shown in Tables 8 and 9. The distributions of scores for each study are shown in Figures 10 and 11.

For Experiment 4, judges were apparently able to discriminate between the English 2 samples and the Arabic ( $\chi^2 [2] = 10.95, p < .01$ ). No differences were found in the ease with which different samples were discriminated ( $\chi^2 [7] = 10.38$ ; Figure 12), and no gender differences were found either (for males  $M = 4.16$ , for females  $M = 4.03$ ;  $t [37] = -0.15$ ).

In Experiment 5, on the other hand, the English 2 samples were seemingly not discriminable from the Chinese ( $\chi^2 [2] = 3.93$ ). However, as in Experiment 4, no differences in discriminability were found among samples ( $\chi^2 [7] = 6.88$ ; Figure 13),

Table 8

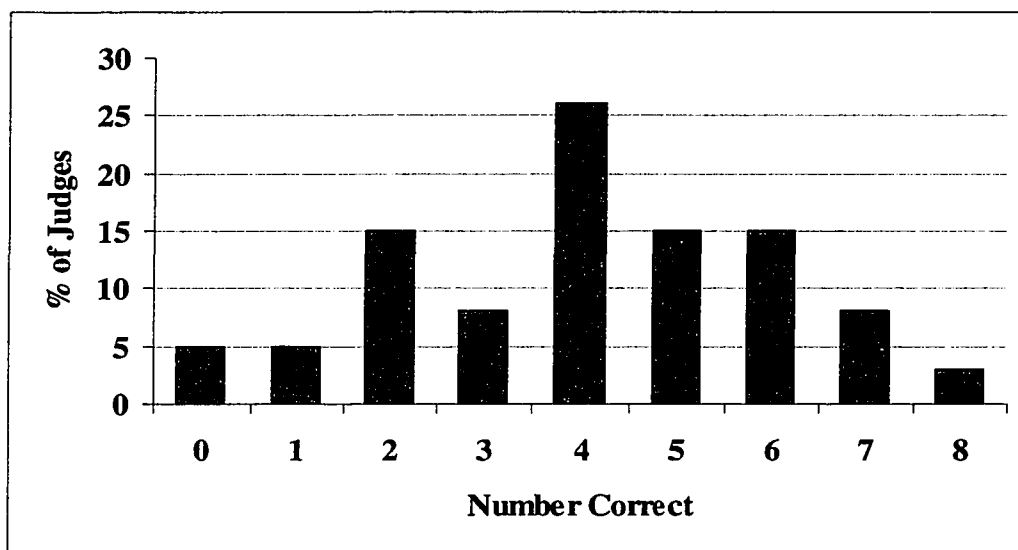
Observed and Expected Frequencies of Scores (Number Correct) on the English 2-Arabic Discrimination Task ( $N = 39$ )

	0, 1, 7, or 8	2, 3, 5, or 6	4
Observed	8	21	10
Expected	2.74	25.6	10.66

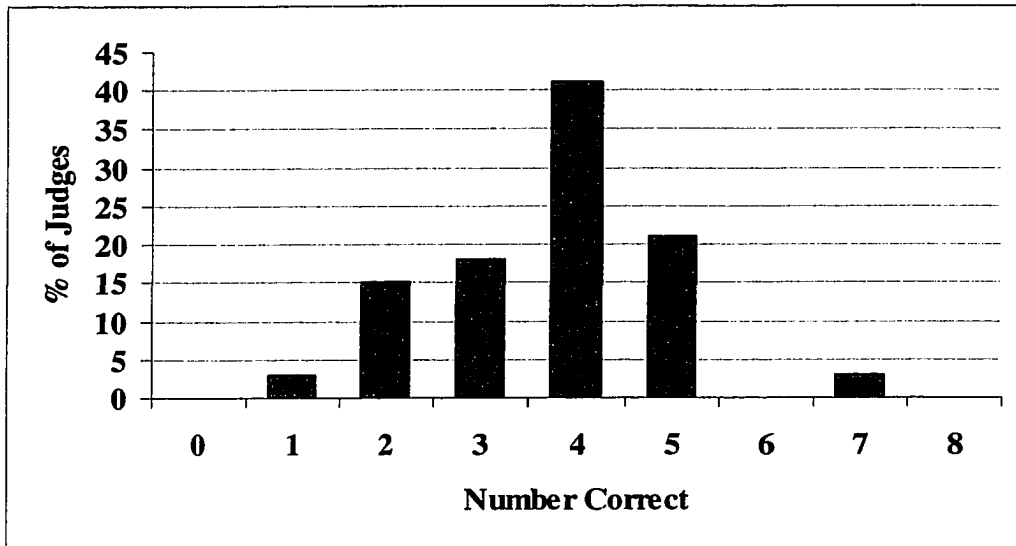
Table 9

Observed and Expected Frequencies of Scores (Number Correct) on the English 2-Chinese Discrimination Task ( $N = 39$ )

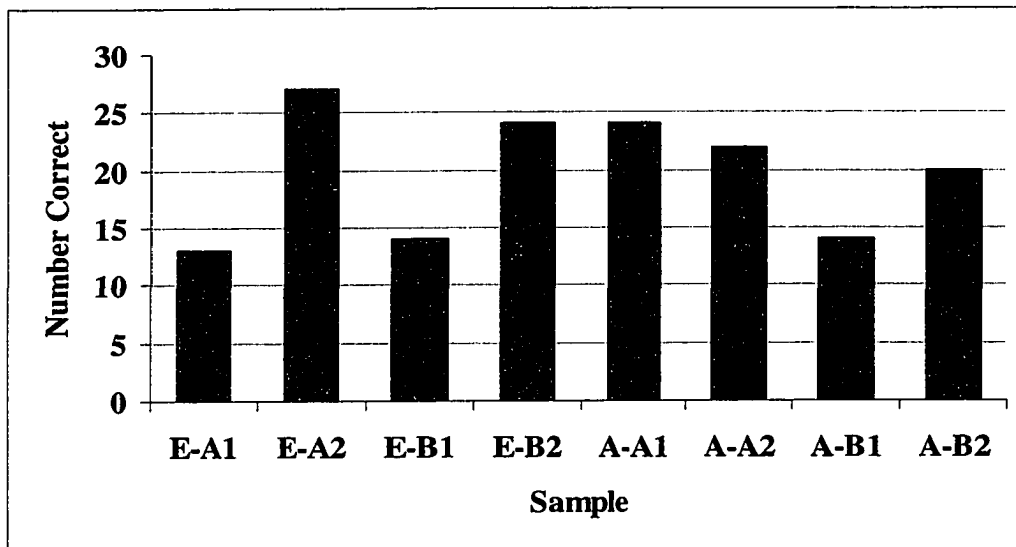
	0, 1, 7, or 8	2, 3, 5, or 6	4
Observed	3	20	16
Expected	2.74	25.6	10.66



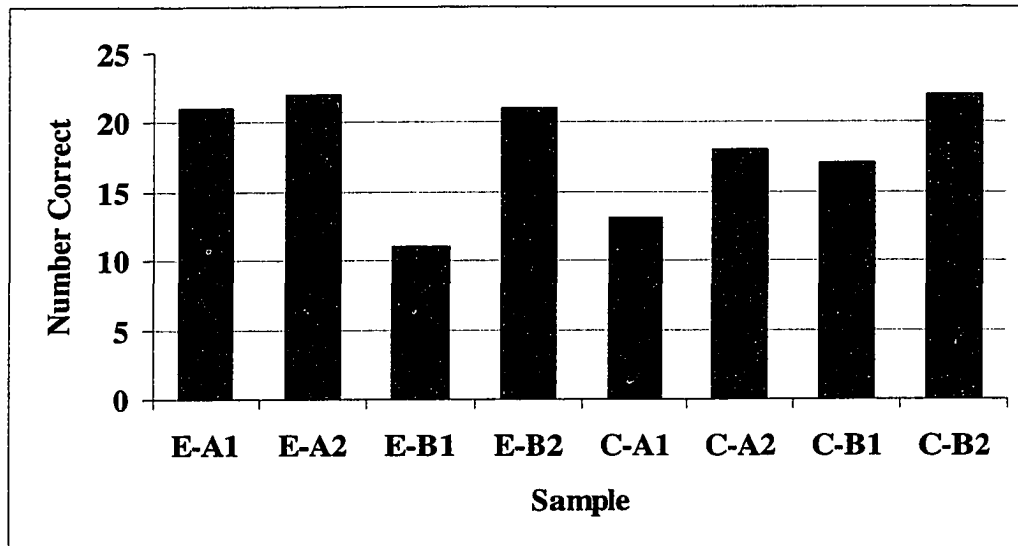
**Figure 10.** Ability of judges to discriminate between English 2 and Arabic babbling samples ( $N = 39$ ).



**Figure 11.** Ability of judges to discriminate between English 2 and Chinese babbling samples ( $N = 39$ ).



**Figure 12.** Number of times the origin of each sample in Experiment 4 was correctly identified ( $N = 39$ ).



**Figure 13.** Number of times the origin of each sample in Experiment 5 was correctly identified ( $N = 39$ ).

and no gender differences were found either (for males  $M = 3.78$ , for females  $M = 3.68$ ;  $t [37] = 0.08$ ). No significant differences were found between language groups in the mean number of consonants (for the English 2-Arabic contrast,  $t [6] = 0.44$ ; for the English 2-Chinese contrast,  $t [6] = -0.73$ ; see Appendix B).

### Discussion

The surprise here was that the samples in Experiment 4 were discriminable, despite the fact that the difference in the number of consonants was not statistically significant (Table 10). The difference in the number of consonants in the English 2 samples in contrast with the Chinese samples was also not significant (Table 11), but, in this case, the samples were not discriminated. This latter result makes sense if number of consonants is the crucial factor in discriminating the samples; the result in Experiment 4 does not make sense, in this context. However, since the difference in Experiment 4 was between *nothing* ( $M = 0$  for the Arabic samples) and *something* ( $M = 0.75$  for the English 2), perhaps the difference was great enough to account for discrimination even though it was not statistically significant.

To shed further light on this question, the next study contrasts samples that are significantly different from each other in number of consonants and should, therefore, be discriminable from each other, if, indeed, the quantity of consonants is the decisive factor.

Table 10

Number of Consonants in Each English 2 and Arabic Babbling Sample--Experiment 4

	E2-A1	E2-A2	E2-B1	E2-B2	A-A1	A-A2	A-B1	A-B2
#	1	1	0	1	0	0	0	0

Table 11

Number of Consonants in Each English 2 and Chinese Babbling Sample--Experiment 5

	E2-A1	E2-A2	E2-B1	E2-B2	C-A1	C-A2	C-B1	C-B2
#	1	1	0	1	5	2	0	1

## CHAPTER 6

### Method & Results

#### Experiment 6: English 1-French Discrimination Task

In this study, the original English samples (English 1) were paired with French babbling samples from which they differed significantly in terms of number of consonants. The prediction was that judges would be able to discriminate the two sets of samples, if consonant quantity was, in fact, what judges were relying upon to make such discriminations.

#### Participants

Ninety-nine judges (23 males, 76 females) were recruited through the participant pool at a local college. Their mean age was 25.5 years (range = 17-60), and all reported having no hearing problems; 14 had children, 46 had studied music, 16 were born outside of the United States, and one had some phonetic training. Thirty-two of these had either studied French or spoke French as a first language.

#### Materials

Samples from three language groups were used for this study: all four original English (English 1) samples; four French samples (which will be described below); and two of the Arabic samples used in Experiments 1 and 4 (samples A-A2 and A-B2, selected at random). The two Arabic samples were inserted randomly into each of the three original orders. The purpose of this change was to attempt to measure attention to the task. The concern was that, since the judges were compelled to participate in

research as a course requirement, they might not give the task their full attention. As the Arabic samples had been quite easy overall for judges to identify as non-English, the thought was that anyone who did not correctly identify either of the two Arabic samples that were inserted into the English 1-French tapes might not be making a good effort and should therefore be dropped from the study.

The French samples, which were once again selected from tapes used by Boysson-Bardies et al. (1984), were chosen at random according to the criteria used in the previous studies. In this study, however, four 8-month-old infants were available, not two as with the Arabic and Chinese tapes. The four samples, therefore, ended up being from four different infants. Three orders (see Appendix A) were once again used, instead of only one as in Experiment 4 and 5.

### Procedure

All testing was done individually in a quiet room using a portable tape player with mini-headphones (Panasonic RQ-P40 [ $n = 67$ ], Sony CFD-540 with Radio Shack Nova 41 headphones [ $n = 19$ ], Aiwa HS-T19 with Sony MDR-013 headphones [ $n = 11$ ], or Sony WM-EX102 [ $n = 2$ ]). Thirty-eight judges heard tape A; 34 heard tape B; and 27 heard tape C.

### Results

The difference between the mean score of the total group ( $N = 99$ ,  $M = 4.05$ ) and the mean of those judges who incorrectly identified both Arabic samples ( $n = 12$ ,  $M = 3.42$ ) did not turn out to be significant ( $t [97] = 0.5$ ). Therefore, the scores of this

latter group are included in the analyses.

Tables 12, 13, and 14 show the observed and expected frequencies of scores for three groups: the total number of judges in this study ( $N = 99$ ); the non-French-exposed judges only ( $n = 48$ ); and the French-exposed judges only ( $n = 32$ ). The distributions for each of these groups are shown graphically in Figures 14, 15, and 16. For none of the groups was there evidence of discrimination:  $\chi^2 (2) = 4.62$  for the total group;  $\chi^2 (2) = 0.89$  for the English-speaking only group; and  $\chi^2 (2) = 2.27$  for the French-exposed group.

Neither order (order A  $M = 3.82$ , order B  $M = 4.06$ , order C  $M = 4.07$ ;  $F [2, 96] = 0.27$ ; Figure 17) nor sex (for males  $M = 4.17$ , for females  $M = 3.9$ ;  $t [97] = 0.71$ ) made a difference in this study; there was an indication, however, that samples differed with regard to how easy they were to correctly categorize ( $\chi^2 [7] = 24.15$ ,  $p < .005$ ; Figure 18). This difference seems to have been accounted for by two samples: E-A1, which was correctly identified by 73 of the 99 judges, and F-A, which was correctly identified by only 29; the mean number of judges correctly identifying any one sample was 49. The two sample sets differed significantly from each other in terms of mean number of consonants ( $t [6] = 3.49$ ,  $p < .005$ , two-tailed; see Appendix B).

### Discussion

The results of this experiment ran contrary to expectations, in several particulars. First, the apparent lack of discrimination between the English and French samples was surprising given that the difference between the mean number of

Table 12

Observed and Expected Frequencies of Scores (Number Correct) on the English-French Discrimination Task: Total Group ( $N = 99$ )

	0, 1, 7, or 8	2, 3, 5, or 6	4
Observed	12	65	22
Expected	6.95	64.98	27.06

Table 13

Observed and Expected Frequencies of Scores (Number Correct) on the English-French Discrimination Task: English-Only Group ( $n = 48$ )

	0, 1, 7, or 8	2, 3, 5, or 6	4
Observed	5	31	12
Expected	3.37	31.51	13.12

Table 14

Observed and Expected Frequencies of Scores (Number Correct) on the English-French Discrimination Task: French-Exposed Group ( $n = 32$ )

	0, 1, 7, or 8	2, 3, 5, or 6	4
Observed	4	22	6
Expected	2.25	21	8.75

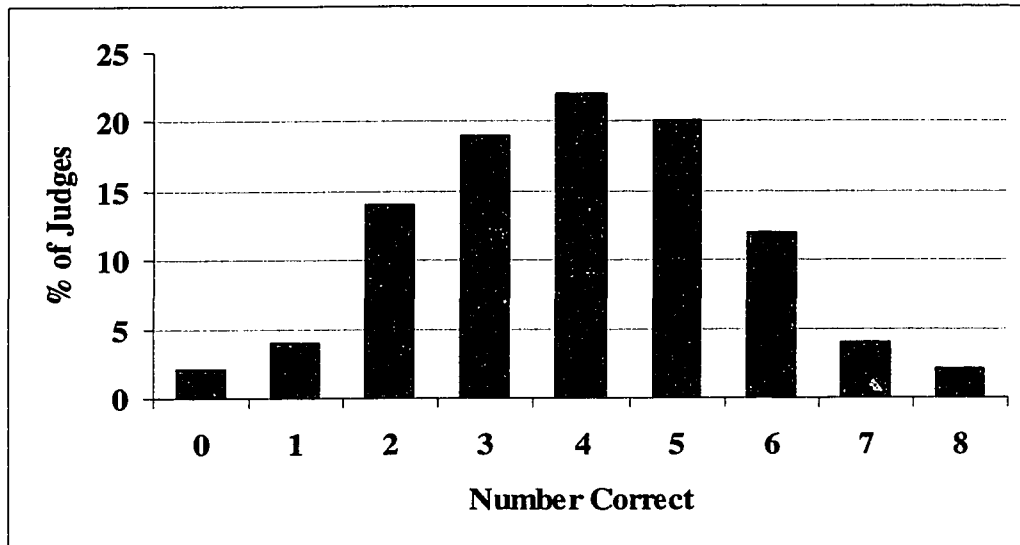
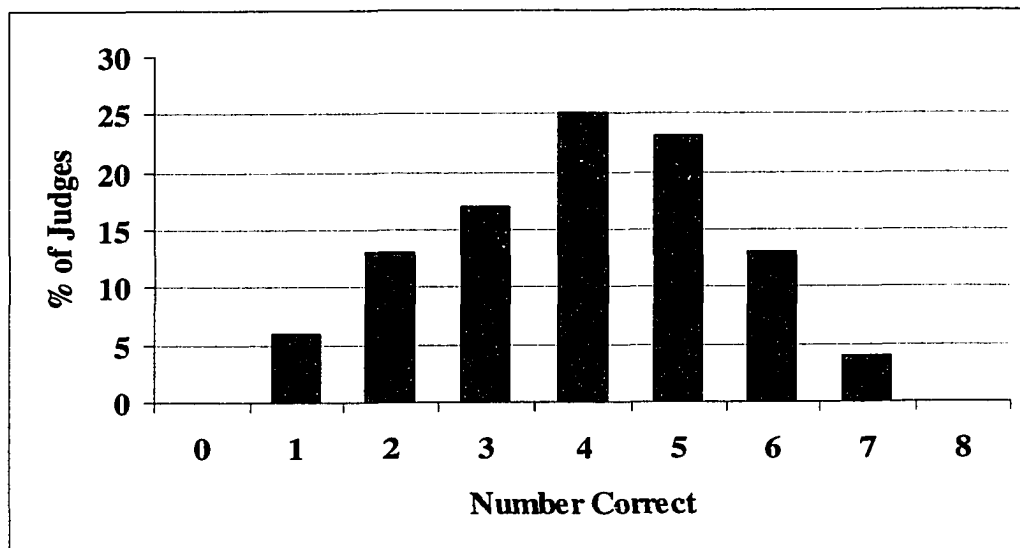
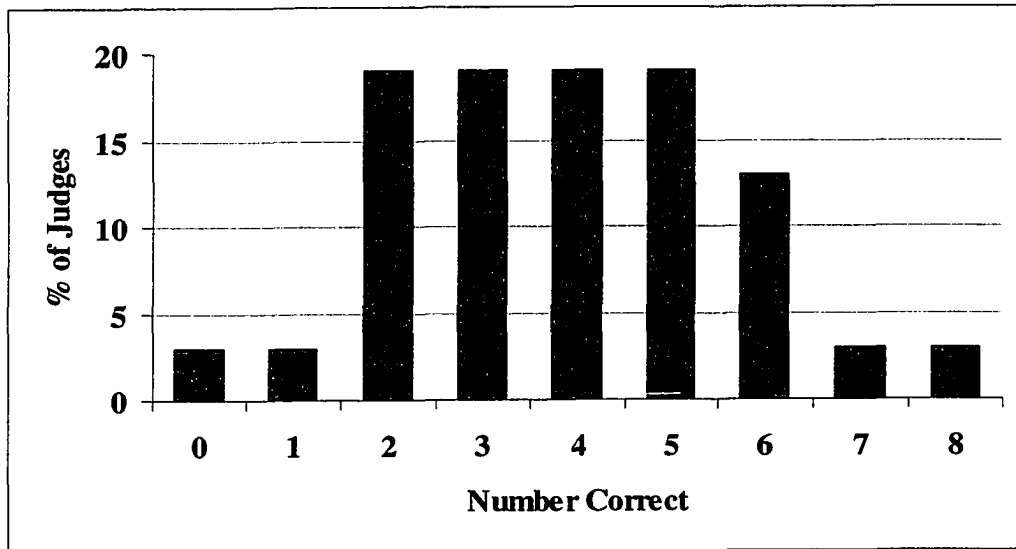


Figure 14. Ability of judges to discriminate between English and French babbling samples ( $N = 99$ ).



**Figure 15.** Ability of judges to discriminate between English and French babbling samples: English-only group ( $n = 48$ ).



**Figure 16.** Ability of judges to discriminate between English and French babbling samples: French-exposed group ( $n = 32$ ).

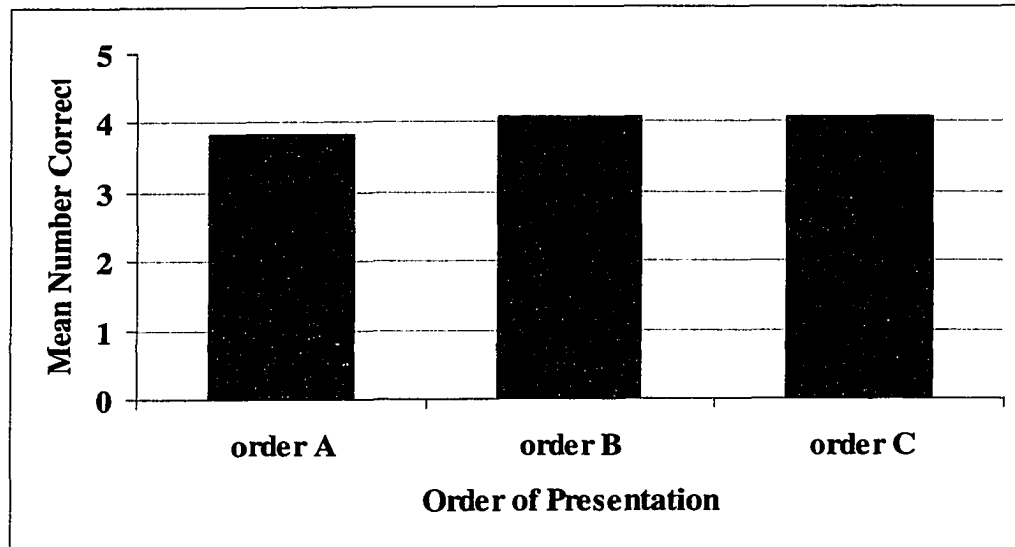
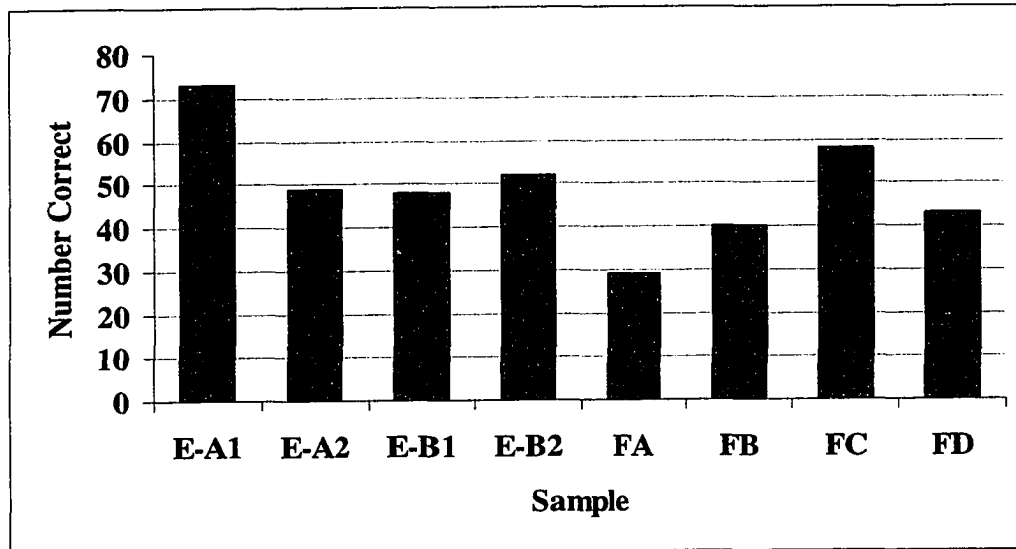


Figure 17. Mean number correct by order of presentation in Experiment 6.



**Figure 18.** Number of times the origin of each sample in Experiment 6 was correctly identified ( $N = 99$ ).

consonants in the samples was statistically significant (Table 15). Second, not all samples were discriminated with the same degree of ease. Specifically, sample E-A1 was much easier to identify than were the others in this study and sample F-A much harder. Several things make these findings especially interesting. Sample E-A1 had been used in three earlier studies. As Table 16 shows, in each of those previous studies, E-A1 had been correctly identified by a proportion of judges that did not differ by a very large margin from the mean proportion correct. In Experiment 6, however, the difference between the proportion of judges who correctly identified E-A1 and the mean proportion of judges correctly identifying any other given sample was quite large. The proportion of judges correctly identifying sample F-A, on the other hand, was quite small--almost as far below the mean as E-A1 was above it. This might seem to indicate that samples with six consonants--which both E-A1 and F-A had--were judged by listeners to be English. However, another French sample also contained 6 consonants, and it was not identified more frequently as English than were other French samples.

Taken together, these results do not seem to support the conclusion that number of consonants is the best or only explanation for the discrimination found in Experiments 1 and 2. Indeed, if that were the case, the samples used in this study should have been distinguished from each other. It is curious that E-A1 should be so much easier to categorize in this context--contrasted with French--than in other contexts. This implies that the samples do in fact contain language-specific cues that may aid in discrimination.

Table 15

Number of Consonants in Each English and French Babbling Sample--Experiment 56

	E-A1	E-A2	E-B1	E-B2	F-A	F-B	F-C	F-D
#	6	8	15	12	6	5	0	6

Table 16

Proportion of Correct Identifications of Sample E-A1 Across Experiments 1, 2, 3, and 6 (N = 99)

	proportion correct identifications of E-A1	mean proportion correct identifications of all samples	difference
Experiment 1	54	62	8
Experiment 2	66	60	6
Experiment 3	52	67	15
Experiment 6	74	49	25

In the final two studies, an attempt was made to eliminate some possible cues in an effort to discover which ones might be responsible for discrimination when it does occur.

## CHAPTER 7

### Method & Results

#### Experiment 7 : Low-Pass-Filtered Babbling Discrimination Task

To see if specific consonants played a role in the discrimination of the English and non-English infants' vocalizations, low-pass-filtered versions of the English 1 and Chinese samples--which had been shown to be discriminable in their unfiltered forms in Experiment 2--were paired in this study. Low-pass filtering at 400 Hertz eliminates most of the information necessary to identify a specific consonant (Kent & Read, 1992), although the perception of a consonantal element remains.

#### Participants

Forty-eight subjects (6 males, 42 females) were recruited through the participant pool of a local college. Their mean age was 19.7 (range = 17-35), and all reported having no hearing problems. Thirteen had children, 17 had studied music, seven were born outside of the United States, and three had some phonetic training. All were native English-speakers, and none had studied Chinese as a second language.

#### Materials

The three tapes used in Experiment 2 were low-pass filtered at 400 Hz using a Stanford Research Systems dual channel filter (Model SR650), with an attenuation slope of 115dB/octave.

#### Procedure

All testing was done individually in a quiet room using a portable tape player

with mini-headphones (Panasonic RQ-P40 [ $n = 37$ ] or Sony WM-EX102 [ $n = 11$ ]). Six judges heard tape A; 31 heard tape B; and 11 heard tape C. Initially, the plan was to use only one order--order B--as in Experiments 4 and 5. Later, orders A and C were added, which is why the groups are of such unequal numbers.

### Results

Many more judges scored 0, 1, 7, or 8 and far fewer 2-6 than would be expected by chance ( $\chi^2 [2] = 68.38, p < .001$ ; Table 17). The distribution this time looked a little different from those in other studies where the two language groups had been discriminated, however (Figure 19): In this study, equal numbers of judges scored at the two extremes (9 judges scored 0-1 and 9 scored 7-8), in contrast to previous studies where scores had clustered toward the higher end. There were no differences by sample ( $\chi^2 [7] = 4.22$ ; Figure 20), by order (order A  $M = 5.66$ , order B  $M = 3.48$ , order C  $M = 4.72$ ;  $F [2, 45] = 2.72$ ; Figure 21), or by sex (for males  $M = 4.83$ , for females  $M = 3.93$ ;  $t [46] = 0.85$ ).

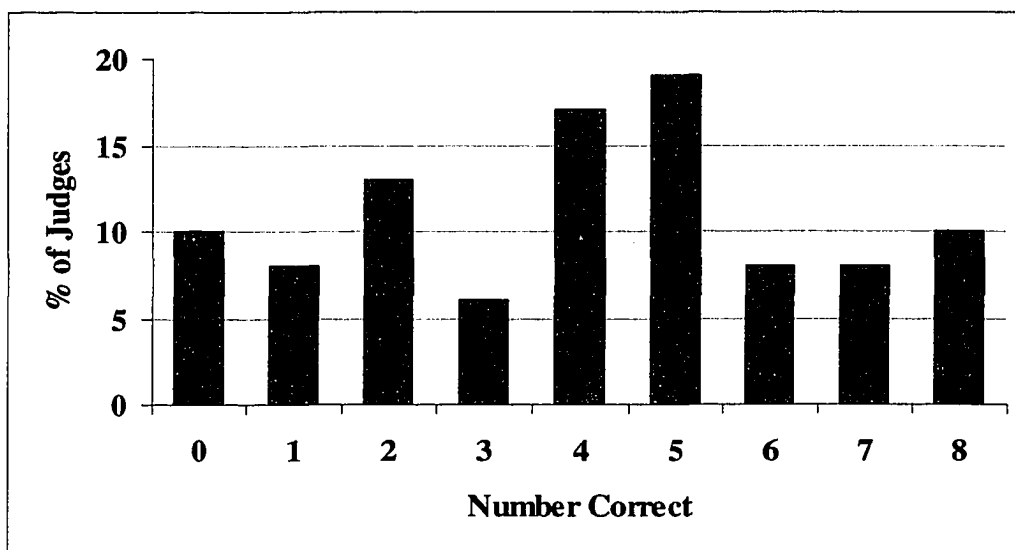
### Discussion

Despite the filtering, the English 1 and Chinese samples were still discriminated. Indeed, filtering the samples does not seem to have made the task appreciably harder than was the original task:  $\chi^2 (2) = 67.25, p < .001$ , for the full-spectrum samples used in Experiment 2; and  $\chi^2 (2) = 68.38, p < .001$ , for the filtered samples in Experiment 7. However, while discrimination overall was as good as in Experiment 2, labeling was inferior: In Experiment 2, 56 % scored above chance, while in this experiment, only

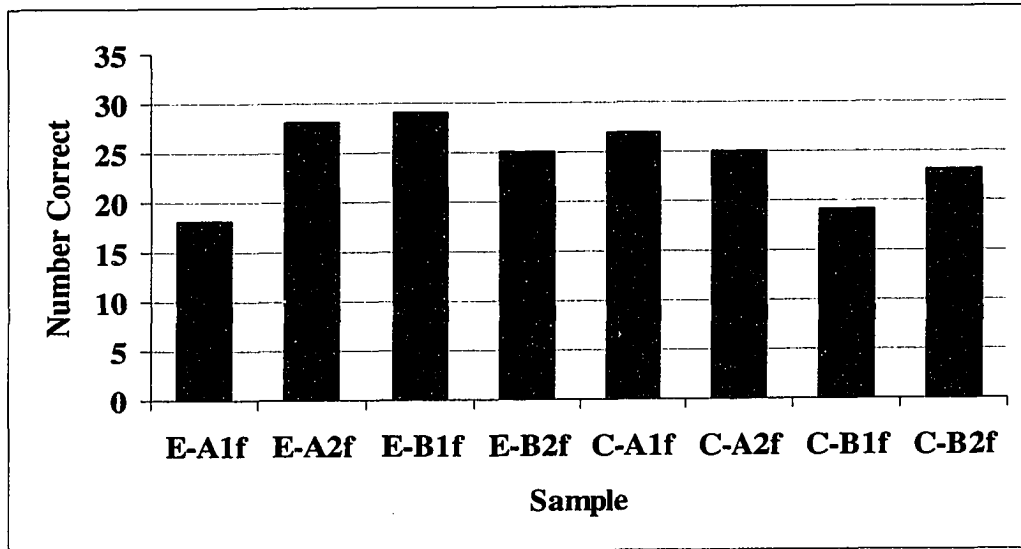
Table 17

Observed and Expected Frequencies of Scores (Number Correct) on the Low-Pass-Filtered Babbling Discrimination Task ( $N = 48$ )

	0, 1, 7, or 8	2, 3, 5, or 6	4
Observed	18	22	8
Expected	3.37	31.5	13.12



**Figure 19.** Ability of judges to discriminate between filtered English and filtered Chinese babbling samples ( $N = 48$ ).



**Figure 20.** Number of times the origin of each sample in Experiment 7 was correctly identified ( $N = 48$ ).

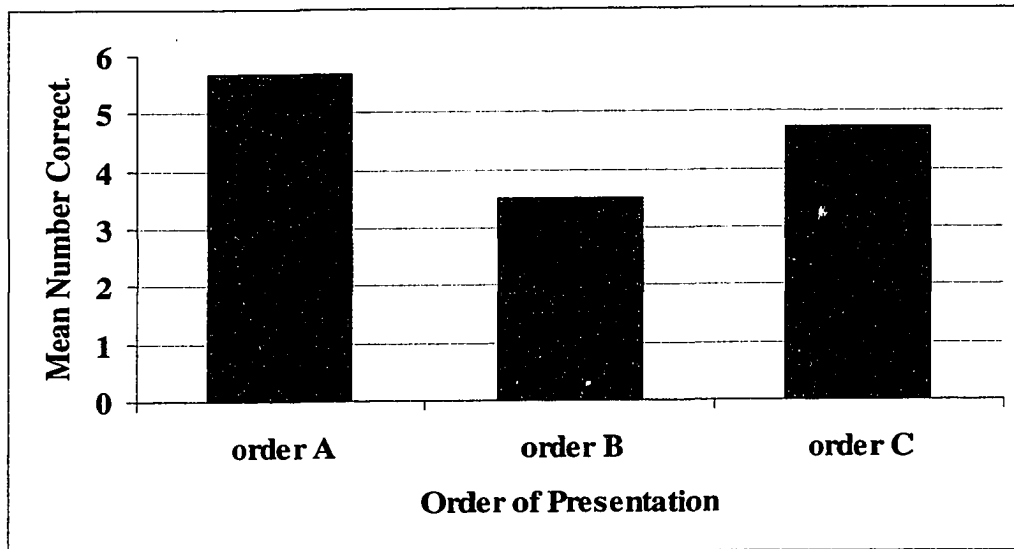


Figure 21. Mean number correct by order of presentation in Experiment 7.

46% did. The relationship between speech type (i.e., filtered or not) and score on the task was not statistically significant, however ( $\chi^2 [8] = 8.14$ ).

It seems that specific consonantal information is not essential for discrimination of the vocalizations of infants from different language backgrounds--or, at least, not for this particular pairing. These results do not rule out, however, the possibility that absolute number of consonants plays a significant role in judges' ability to discriminate some of the sample sets used in these studies. Language-specific prosodic differences between the vocalizations of infants from different language environments are also not ruled out as an explanation.

The final experiment is designed to look at what happens to the discriminability of samples when prosody is disrupted but the actual sounds, as well as the number of consonants, remain unchanged.

## CHAPTER 8

### Method & Results

#### Experiment 8: Scrambled Babbling-Natural Babbling Discrimination Task

In this eighth study, English 1 samples that had been scrambled (i.e., the sounds within the samples were rearranged) were contrasted with the original English 1 samples. The purpose of this manipulation was to see if disrupting the language-specific prosody of the speech, without altering its phonetic content, would make such samples discriminable from unaltered samples. If so, a prosodic-differences hypothesis would be supported, rather than a number-of-consonants one.

#### Participants

Fifty-four students (14 males, 40 females), with a mean age of 19.5 (range = 17-26), were recruited through the participant pool of a local college. All reported normal hearing and English as their first language. Four had children, 22 had studied music, 16 were born outside of the United States, and one had some phonetic training.

#### Materials

The four English 1 samples were used again in this study, in two forms: in their natural form and in a scrambled form. The scrambled samples were created in a computer program called WaveStudio 2.0 on an IBM Aptiva 530 (Model 2144-66P) equipped with a Sound Blaster 16 audio card. In this program, it is possible to view the wave form of a sound file and cut and paste segments to different locations within the sample.

The four English 1 samples were first digitized at 22 kHz, using this same program. Segments suitable for rearranging were identified visually on the displayed wave form and then played out through the computer's speakers to make sure that the cuts didn't come in the middle of a sound. The segments were then reordered completely randomly within the same time envelope so that the overall length of the scrambled samples remained the same as the natural samples from which they derived. Finally, the natural and scrambled samples were recorded from the computer through a Marantz tape recorder (Model PMD22) onto an audiotape, in the same three orders that were used in all previous studies (see Appendix A).

#### Procedure

All testing was done individually in a quiet room using a portable tape player with mini-headphones (Sony WM-EX102 [ $n = 52$ ] or Panasonic RQ-P40 [ $n = 37$ ]). Twenty judges heard tape A; 12 heard tape B; and 22 heard tape C. The inequality of these groups resulted from a disproportionate number of participants who heard tape B being disqualified because of bilingualism. As in Experiment 3, instructions to participants remained unchanged, even though in this experiment, as in that one, all samples, in fact, derived from American infants.

#### Results

Identifying English 1 as "English" and Scrambled English 1 as "not English" were considered to be the correct responses in this study. Observed and expected frequencies of scores, shown in Table 18, were not significantly different:  $\chi^2(2) = 2.28$ .

Table 18

Observed and Expected Frequencies of Scores (Number Correct) on the Scrambled Babbling-Natural Babbling Discrimination Task ( $N = 54$ )

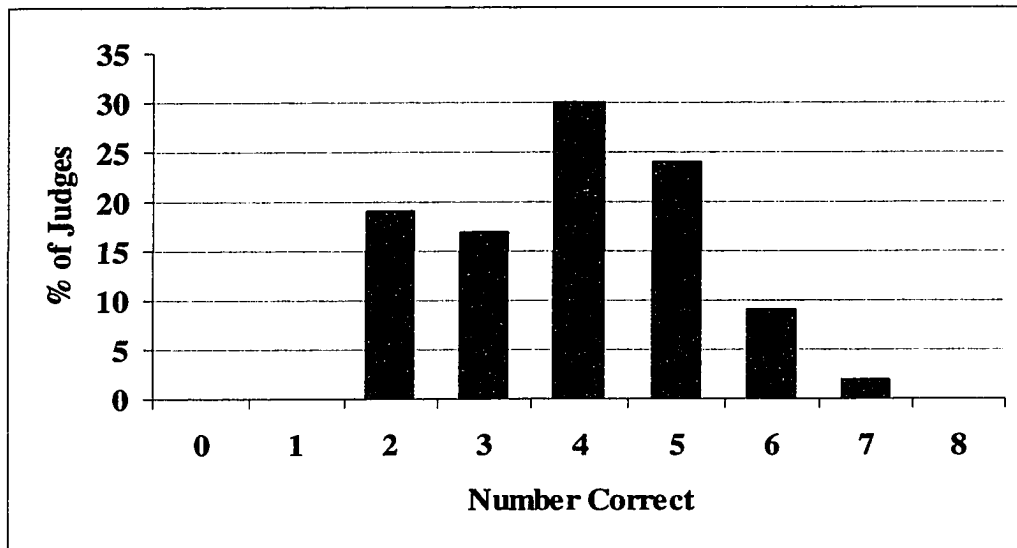
	0, 1, 7, or 8	2, 3, 5, or 6	4
Observed	1	37	16
Expected	3.79	35.45	14.76

The distribution (Figure 22) was, in fact, very close to a normal one, the only difference being that somewhat fewer judges were able to discriminate--as defined by a score of 0-1 or 7-8--between the scrambled and natural samples than would be expected by chance.

Order was not a significant factor (order A  $M = 4.2$ , order B  $M = 3.58$ , order C  $M = 3.9$ ;  $F [2, 51] = 0.81$ ; Figure 23), nor was sex (for males  $M = 4.21$ , for females  $M = 3.85$ ;  $t [52] = 0.92$ ). As in Experiment 7, however, there was a difference in this experiment in the discriminability of samples ( $\chi^2 [7] = 17.79$ ,  $p < .025$ , two-tailed; Figure 24). Again, E-A1 appears to be primarily responsible for the difference, with 40 of the 54 judges correctly identifying it as English, when the mean for all the samples was only 26.6. The scrambled counterpart to this sample, E-A1-S, with only 15 judges correctly identifying it, may also have contributed to the statistically significant difference found.

### Discussion

Despite the reordering of the sounds within the samples, the English 1 and Scrambled English 1 samples were not discriminated. If judges rely upon prosody alone as their guidepost, these two sample sets should have been discriminated. The fact that they were not indicates that discrimination, when it does occur, may be made on some basis other than prosodic differences. Alternatively, prosody may not have been sufficiently disrupted in this instance to affect discrimination one way or the other, a point that will be returned to in the General Discussion. Furthermore, a difference in the



**Figure 22.** Ability of judges to discriminate between English and scrambled English babbling samples ( $N = 54$ ).

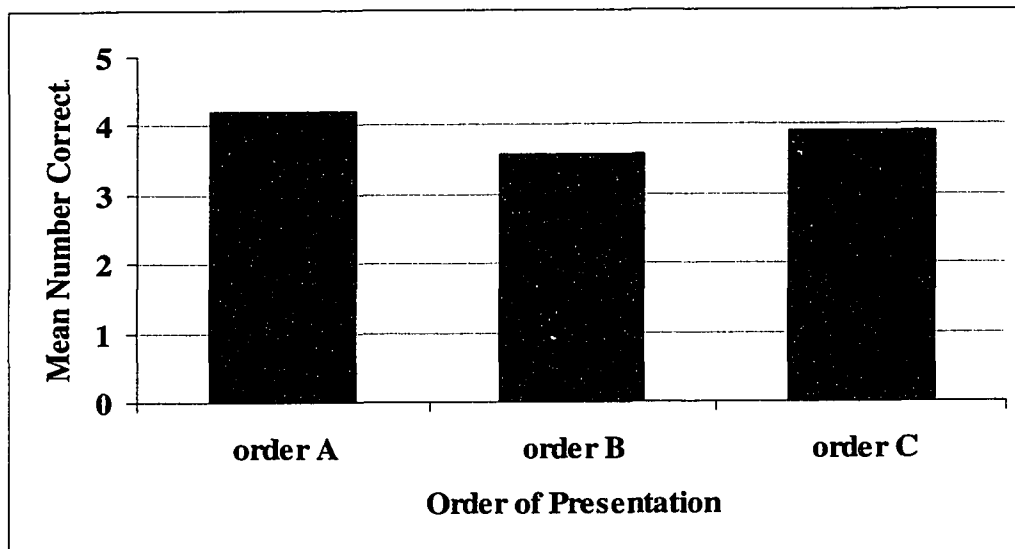
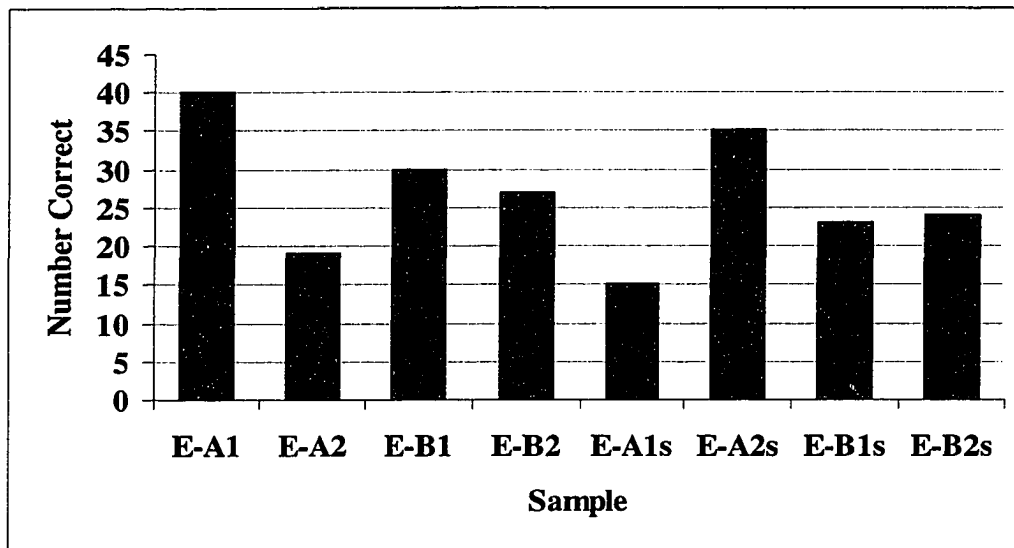


Figure 23. Mean number correct by order of presentation in Experiment 8.



**Figure 24.** Number of times the origin of each sample in Experiment 8 was “correctly”<sup>1</sup> identified ( $N = 48$ ).

---

<sup>1</sup>“Correctly,” in this case, means identifying the unscrambled samples as “English” and the scrambled samples as “not English.”

number of consonants has still not been ruled out as a possible basis for discrimination, though the recurring finding that some samples are discriminated more easily in some contexts than in others continues to create a problem for a strict quantity-of-consonants explanation.

## CHAPTER 9

### General Discussion

Eight studies were undertaken to further understanding of the course of early prosodic and phonetic development. Specifically, these studies were designed to extend Boysson-Bardies et al.'s (1984) work and to discover what might be the basis or bases for adult listeners' ability to discriminate between babbling samples of infants from their own and from another language group. The results of the present studies will first be summarized individually and will then be discussed collectively to see what support, if any, the data offer for several different possible explanations for the discriminability of the vocalization samples. In a final section, unanswered questions will be considered, along with suggested future directions.

#### Summary of the Results of Experiments 1 through 8

*Experiment 1—English 1 contrasted with Arabic:* English 1 samples and Arabic samples were easily discriminated by the judges. Moreover, they got the labels right as well: More than half of the judges correctly identified 75% or more of the samples (i.e., scored 6 or better on the task). There was no indication that some samples were easier/harder to identify than others, but an order effect and a sex difference were found, possibly because of unequal groups. The number of consonants in the samples differed significantly by language, which could account for the discrimination.

Experiment 2–English 1 contrasted with Chinese: English 1 and Chinese samples were also discriminated, but with only about half the listeners achieving a score of 6 or above. As equal numbers of males and females were tested this time and nearly equal numbers of participants were tested on each order, neither an order effect nor a sex difference was found in this experiment. Again, samples did not differ with regard to ease of discrimination. However, since consonant quantity once again varied as a function of language group, it remained possible that discrimination was made on that basis alone.

Experiment 3–English 1 contrasted with English 2: Judges were able to distinguish between one set of English samples (English 1) and another set of samples from same-age English-environment infants (English 2) whose babbling contained significantly fewer consonants than did the first group. Interestingly, they reliably labeled the English 1 samples as “English” and the English 2 samples as “non-English.” No difference was found in the discriminability of the samples, nor was there an order effect. A sex difference was found, but with only 4 males tested out of a total of 32 participants, this result was suspect. Since the infants were all the same age and were all from English-speaking homes, the number of consonants in the samples seems to be the most likely explanation for their discriminability.

Experiment 4–English 2 contrasted with Arabic and Experiment 5–English 2 contrasted with Chinese: In both cases, the number of consonants in the samples did not differ significantly as a function of language group. Nevertheless,

whereas English 2 and Chinese were not discriminated, English 2 and Arabic were. Only one order was used in these studies; no differences were found by sex or by sample.

*Experiment 6–English 1 contrasted with French:* These samples were not discriminated, despite the fact that they differed significantly from each other in terms of number of consonants. Neither order nor sex differences were found, but this time differences in the discriminability of the samples were found: sample E-A 1, which had been used in three other studies, was much easier to discriminate in this context, while F-A seemed to be harder than the other samples to discriminate.

*Experiment 7–filtered English 1 contrasted with filtered Chinese:* Low-pass-filtered versions of the English 1 and Chinese samples were discriminated just as easily as were the full-spectrum originals. Labeling, however, was inferior, with 56% scoring above chance with the full-spectrum samples and only 46% doing so with the filtered versions, but the relationship between speech type and score on the task was not statistically significant. No order, sample, or sex differences were found.

*Experiment 8–English 1 contrasted with scrambled English 1:* Scrambled and natural versions of the English 1 samples were not discriminated from each other. Neither sex nor order was a factor. Once again, however, sample E-A 1 was easier to categorize in this context than it had been in some others, while its scrambled counterpart seemed to be particularly difficult to categorize.

In many but not all instances, then, adult English-speaking judges do appear to

be able to distinguish between 8-month-old infants from different language backgrounds. Two exceptions found here were the English 1-French pairing and the English 2-Chinese pairing. Background factors such as age, sex, musical training, phonetic training, participant's country of origin, equipment used, headphone position, having children or not, and number of times the tapes were heard had no detectable effect on performance. Order did not appear to have an effect on the task, except in Experiment 1 where group sizes were very lopsided and order was confounded with phonetic training. In general, samples did not differ with regard to the ease or difficulty with which they were categorized. The few exceptions to this rule will be discussed in later sections.

But while the question, "Is it possible to distinguish between the babbling of infants from varying language backgrounds?" might be answered with a qualified "yes"--notwithstanding the two deviations mentioned above--the question of just what the basis for the discrimination might be remains unanswered. Is the ability to discriminate that has been shown here due to ambient-language effects on infant vocalizations at this early stage? Or is there some other, possibly nonlinguistically-based, explanation for these results? And if ambient language is playing a role in the discriminability of babbling at this age, is its effect manifesting itself in the prosodic or the phonetic aspects of the infants' speech? The data from the eight experiments conducted will be considered below with reference to four different possible bases for discrimination.

### Quantity of Consonants

The first explanation I considered was that the judges were discriminating samples based solely on the number of consonants each contained. (Table 19 shows the outcomes of all eight experiments, as well as which pairings differed significantly with regard to mean consonants per sample.)

For Experiments 1 and 2, this explanation made perfect sense. In Experiment 1, English 1 samples, with a mean consonant count per sample of 10.25, were paired with Arabic samples, with a mean consonant count of 0. In Experiment 2, English 1, with its many consonants, was contrasted with Chinese, with a mean consonant count of only 2 per sample. These differences were statistically significant, with the protected  $t$  for the English-Arabic pairing equal to 5.86,  $p < .001$ , two-tailed, and the protected  $t$  for the English-Chinese pairing 4.39,  $p < .01$ , two-tailed.

This inspired me to contrast English 1 with new English samples from same-age infants (English 2), which is what was done in Experiment 3. Here, again, the difference in the quantity of consonants in English 1 and English 2 was significant ( $t[7] = 5.86$ ,  $p < .001$ , two-tailed), and the samples were discriminated. This further bolstered the notion that the discriminability of the samples was not due to ambient-language effects; rather, it was due to the number of consonants in the samples. Of course, it was possible that the difference in the number of consonants in the samples was, in fact, an ambient-language effect. There is, however, no evidence supporting such a difference in babbling between infants of varying language backgrounds. Such an explanation,

Table 19

Summary of Results: Experiments 1 through 8

<b>Languages Contrasted</b>	<b>Significant Difference in Consonants?</b>	<b>Discriminated?</b>
Exp. 1: English 1-Arabic	yes	yes
Exp. 2: English 1-Chinese	yes	yes
Exp. 3: English 1-English 2	yes	yes
Exp. 4: English 2-Arabic	no	yes
Exp. 5: English 2-Chinese	no	no
Exp. 6: English 1-French	yes	no
Exp. 7: Filtered English 1- Filtered Chinese	yes	yes
Exp. 8: English 1- Scrambled English 1	no	no

moreover, would not serve for the results of Experiment 3, where all infants were growing up in English-speaking households. The more likely explanation, therefore, was that the differences were due to sampling and not to language background.

The results of Experiments 4 and 5 did not entirely support this interpretation, though. The quantity-of-consonants explanation would predict that neither the English 2-Arabic nor the English 2-Chinese pairing would be discriminated, as neither the Arabic nor the Chinese sample sets differed significantly from English 2 in terms of consonant number ( $t[7] = 0$  for the English 2-Arabic pair;  $t[7] = 1.48$  for the English 2-Chinese pair). While this prediction was borne out for the English 2-Chinese contrast, it was not for the English 2-Arabic one.

Experiment 6 dealt a further blow to the consonant-quantity model. This time, the model would have predicted that English 1 and French, differing as they did in mean consonants per sample (10.25 for the English; 4.25 for the French;  $t[7] = 2.51$ ,  $p < .05$ , two-tailed), would be discriminable. They were not. In addition, the fact that two samples--one English, one French--differed from the others in terms of the ease with which they were discriminated further undermined the consonant-quantity model.

The final two experiments did not directly address the question of consonant number. In Experiment 7, the low-pass-filtered English 1 and Chinese samples were discriminated just as easily as the unfiltered versions had been. The filtering did nothing to change the number of consonants per sample, although it did remove most of the information necessary to identify the specific consonants. Labeling, on the other hand,

was affected by filtering, with the judges performing better on the unfiltered than on the filtered versions. This result will be discussed again later with reference to other possible explanations. Of interest here, though, is the indication that listeners are doing more than merely counting consonants in coming to a decision regarding the origin of babbling samples. Otherwise, labeling would be expected to be as accurate for the filtered as for the unfiltered samples, as the number of consonantal elements was unchanged.

In Experiment 8, the number of consonants was identical, as indeed were the actual consonants. Had the samples been discriminated, that would have been a powerful argument against the quantity-of consonants model. As the samples were not discriminated, it is not clear what role number of consonants or specific consonants may have played.

In sum, there is evidence that number of consonants per sample may play a role in judges' ability to discriminate sets of babbling samples. But there is evidence as well that this may not be the sole basis upon which discriminations are made. At this stage, it is only possible to say that consonant quantity provides neither a necessary nor a sufficient basis for discrimination, though it may help in some cases. Further studies, which will be discussed in a later section, will be required to resolve this question.

#### Background Noise

Another possible explanation of the results of these experiments rests on the

fact that the samples used were recorded by different people, under varying conditions, using different equipment. Perhaps English 1 and Arabic, for example, were discriminated by the judges not because of the number of consonants and not because of any inherent differences due to ambient-language influences, but because the two sets of samples differed with respect to background noise.

Such a possibility cannot be completely ruled out on the basis of the results from the current studies, although additional studies (to be discussed later) might shed some light on this question. Nevertheless, since differences in background noise are present in all sets--as indeed they exist to some degree between samples within the same language set--the finding that not all language pairings were discriminated seems to argue against such an explanation. Additionally, some differences in recording quality, echoing, etc., between the English 1 and Chinese samples were masked by the filtering done in Experiment 7, yet these samples were still discriminated as easily as before.

#### Ambient Language Influence: Prosody

Boysson-Bardies et al. (1984) believed that the explanation of their findings lay in infants' incorporation of language-specific prosody into their babbling. The results of the present studies are unclear with regard to the possible role of prosody in adult judges' ability to discriminate between the babbling of infants from a variety of language environments.

If prosodic differences are present in the babbling of 8-month-old infants to the degree that discrimination on that basis would be possible, then discrimination would be the expected result in Experiments 1, 2, 4, 5 and 6. However, the English 2-Chinese pairing (Experiment 5) and the English 1-French pairing (Experiment 6) were not discriminated. On the other hand, discrimination would not be the predicted result in Experiment 3, as two sets of samples from 8-month-old American/English-environment infants would not be expected to differ with regard to language-specific prosody. English 1 and English 2 were discriminated, however.

Experiment 7 did not rule out prosody as a possible basis for discrimination, as low-pass filtering exercises its effect primarily on the phonetic, not the prosodic, aspects of the speech signal. But in Experiment 8, the disruption of normal prosody by scrambling the samples did not yield the expected result of discrimination of the samples.

All in all, the support for language-specific prosody as the basis for discrimination is fairly weak at this point. Perhaps it plays a role, but, as with consonant number, it does not seem to be a necessary or sufficient ingredient. Additional studies, described later on, might provide some more definitive answers, however.

#### Ambient Language Influence: Phonetics

By 9-10 months, it has been reported that infants exhibit differences in the frequency of occurrence of particular consonants in their babbling that reflect

differences in the adult target languages (Boysson-Bardies & Vihman, 1991). These 9-10-month-olds were the youngest infants included in this particular study, but it may be that these differences are already in evidence to some extent prior to that age--perhaps even by 8 months, the age of the infants used in the present eight studies. Even if the differences at this earlier stage have not yet reached statistical significance, it may be that they are detectable by adult judges, who can use them to make determinations regarding the language background of the infants to whose vocalizations they are listening.

If this is so, what support is there from the current work for a phonetic basis for the results? The same experiments that caused problems for the prosodic-differences explanation of the results create difficulties for the phonetic-differences explanation. As above, if language-specific phonetic differences were present and discernible in the samples, then discrimination should have been found in Experiments 1, 2, 4, 5, and 6; instead, discrimination was found only in Experiments 1, 2, and 4. Discrimination should not have been found in Experiment 3, as phonetic differences would not be expected between English 1 and English 2; but this pairing was discriminated.

Experiments 7 and 8 do provide some hints of a significant role for phonetic differences, however. In Experiment 8--where no phonetic differences between the sample sets existed--no discrimination was found. This may not offer specific support for a phonetic interpretation, but it does not refute it either. More interesting are the results of Experiment 7. In this study, discrimination was just as good with the filtered

as with the nonfiltered samples. This appears to argue against specific consonantal information as a basis for discrimination. It is important to note, however, that labeling was more accurate with the full-spectrum samples than with the filtered ones. In other words, listeners were better able to correctly identify the origin of the samples when the full-spectrum cues were available than when they had only the low-frequency information available. This implies that, while it may be easy enough for listeners to identify two distinct groups using low-frequency information alone, they need the high-frequency information to correctly categorize the groups. As before, further studies that will be outlined below may be able to clarify the role of high-frequency (consonantal) information in discriminating the babbling of babies from different language backgrounds.

#### How Is Discrimination Possible?

What, in conclusion, appears to be the basis for discrimination of the babbling of infants from different language environments, when discrimination does occur? Is there anything that the infants themselves are doing that aids discrimination? Or is discrimination based solely on such factors—probably artifactual—as consonant number and background noise?

Certainly factors related to sampling—which encompasses both the audiotaping and the selection of the segments of infants' vocalizations for use in the studies—have the potential to convey information to listeners, which they may, in turn, use in making

judgments about the origin of the samples. But just because the information is there, can we be sure that the judges are making use of it? And can we be sure that that is the *only* information the judges are relying upon to make their determinations?

The results of the studies described here do not provide a clear-cut answer to these questions, but they do provide some clues as to whether and how much this nonlinguistic information is being used by listeners to arrive at their decisions. With reference to the first question, it is hard to understand why, since background noise is present in all cases, this information would not be used in some instances—for example, in Experiments 5 and 6, where the pairings were not discriminated—but would be used in others. Similarly, if consonant quantity is critical, how can we explain the results of these same two experiments? It seems that, if these factors—that is, differences in the tape recordings and differences in the mean number of consonants per sample—were indeed the basis upon which the judges made their discriminations, how would it be possible to explain those cases where the two sample sets were not discriminated?

Several other findings seem to conflict with an explanation based solely on consonant number or background noise. First, of all the background factors examined (see Appendix B), only phonetic training appeared to have any influence on task performance. If phonetic training makes a judge better at discriminating between infants who are being exposed to different language environments, then it seems likely that the difference between the samples is in the linguistic realm. However, this argument is not

as strong as it might be for two reasons: First, phonetic training might attune listeners to consonants and background noise, so they might, in fact, be better able to make use of that information than naïve judges would be; and, second, most of the phonetically-trained judges participated in Experiment 1, so it is hard to know whether the training only benefitted them in that particular task or if the advantage would have been more general.

Also problematic for the nonlinguistic-factors explanations are the findings of differences in the ease or difficulty with which some samples are categorized. How can the fact that sample E-A1, for example, is so much easier to discriminate in the context of the English l-French contrast than in any of the other pairings be explained on the basis of consonant number or background noise?

All in all, then, it seems that, in order to explain the findings, it is necessary to assume some kind of ambient-language effect on the infants' vocalizations. An explanation based on prosody or phonetics surely seems essential to understanding the varying discriminability of samples in different contexts. The discovery in contrasting Experiments 2 and 7 that listeners' labeling was affected by the filtering but not their overall ability to discriminate also seems to require an acknowledgment of ambient-language influences—particularly in the phonetic sphere. However, some of the same issues arise in considering a prosodic- or phonetic-factors explanation that arose in relation to the sampling-factors explanations. Again, if we believe that infants' babbling is substantially influenced by the adult target language at this

stage, how do we understand the results of Experiments 4 and 6? And Experiment 3 must also be viewed as an anomaly.

### Future Directions

Although there are strong indications that some aspects of the ambient language are making their way into infants' speech even as early as 8 months of age, there are still a good many unanswered questions and as many possible future studies. Below are just a few of these possibilities that seem to me to represent the next logical steps in this research program.

Collection of new babbling samples: The best way to address questions about the role of acoustic differences between the samples would be to collect new babbling samples under acoustically controlled conditions, followed by a strict sampling procedure.

Filtered English 1 paired with filtered and scrambled English 1: It is possible that the lack of discrimination found in Experiment 8 was due to the fact that not only were there the same number of consonants in both sets, but the consonants were identical, although reordered. Perhaps this overpowered any effect that the reordering of the elements might have had on discrimination. If the samples were low-pass filtered, most of the specific phonetic information would be removed, which might then increase the salience of the prosodic disruption that the scrambling was meant to achieve.

More thoroughly scrambled English 1 paired with English 1: Another approach might be to scramble the samples more thoroughly. In Experiment 8, the individual sounds were reordered within the original envelope. However, *da-da-da* doesn't sound that different when it is rearranged. Perhaps it would be useful to explore other ways to tamper with the prosody of the samples that might have more dramatic effects.

High-pass-filtered versions of English 1 and Chinese: To attempt to rule out background noise as a factor in discrimination, high-pass-filtered versions of the English 1-Chinese pairing (or any other pair that were discriminated) could be prepared. Additionally, if high-pass-filtered samples were discriminated, that would indicate that prosody might not be crucial to discrimination, as the lower frequencies carry most of the prosodic information. Such an outcome would also imply that specific consonants are in fact a determining factor in identifying the language background of an infant even at this early age.

Acoustic analyses of the samples: It would be very informative to perform acoustic analyses of the samples used in these eight studies. Of particular interest would be those cases where the language sets were not discriminated, such as in Experiments 5 and 6.

The studies reported on in chapters 2-8 truly represent only the tip of the iceberg, in that they are suggestive with regard to possible bases for discrimination but do not provide definitive answers. The research program outlined above should begin

to fill in some of the gaps left by Experiments 1 through 8, as well as by the previous research done in this area, leading, perhaps, to some better, more focused questions, if not to firmer conclusions.

## APPENDIX A

### Orders of Presentation of Babbling Samples for Experiment 1

<u>order A</u>	<u>order B</u>	<u>order C</u>
E-A2	A-B1	☛
A-B2	E-B2	E-B1
E-B1	E-A1	A-A2
☛	A-A2	A-B1
A-A2	E-A2	E-A2
E-A1	☛	☛
A-B1	A-B2	A-B2
A-A1	☛	E-A1
E-B2	E-B1	E-B2
☛	A-A1	A-A1


**Note:** Each sample is designated by two letters and a number, as follows: The first letter identifies the language environment of the infant ("E" for English; "A" for Arabic); the second letter identifies which baby the sample came from ("A" for baby A; "B" for baby B); the number identifies the specific sample from that baby ("1" for sample 1"; "2" for sample 2). Thus, sample E-A2 is sample number 2 from English baby A and sample A-B1 is sample number 1 from Arabic baby B.

**Experiment 2:** The same orders were used, with C-A1 used in place of A-A1, C-A2 used in place of A-A2, and so on.

**Experiment 3:** The same orders were used, with E2-A1 taking the place of A-A1, E2-A2 replacing A-A2, and so on.

**Experiments 4& 5:** In Experiment 4, order B was used, with E2 samples replacing E1 samples, and the Arabic samples remaining the same; in Experiment 5, order A was used, with E2 samples replacing and Chinese samples replacing the Arabic, as in Experiment 2.

**Experiment 6:** The same orders were used in this study, with F-A (French sample A) replacing A-A1, F-B replacing A-A2, F-C replacing A-B1, and F-D replacing A-B2. In order A, A-A2 was inserted between the third and fourth samples and A-

B2 was inserted following the eighth sample. In order B, A-A2 followed the fifth sample and A-B2 followed the sixth sample. In order C, A-A2 was inserted at the beginning and A-B2 was inserted between the fourth and fifth samples. The symbol  above indicates where an insertion occurred.

Experiment 7: The same orders were used as in Experiment 2.

Experiment 8: The same orders were used, with E-A1-S ("S" designates a scrambled sample) used instead of A-A1, E-A2-S used instead of A-A2, and so on.

### Instructions to Participants

On this tape are recorded 8 short babbling samples from 8-month-old infants. Each sample is approximately 10 seconds in length, and there are 10-second silences between samples. Some of these infants are growing up in households where English is spoken by their caregivers, and some are growing up in households where some other language is spoken. Your task is to attempt to identify which are the infants who are growing up in the English-environment households and which are not.

Please listen to the tape once all the way through without writing anything on the response sheet. You can then rewind the tape and listen as many more times as you like. When you are ready, simply write "English" or "not English" for each sample. You are not expected to try to identify what the language background of the non-English infants is. Please do not leave any blanks; you must write something for each sample. You will also need to answer the questions printed at the top of the response sheet and, when you are through, write down how many times you listened to the tape.

(The form the participants signed was on CUNY Graduate Center letterhead and was approved by the Committee on the Protection of Human Subjects at the CUNY Graduate Center.)

### CONSENT FORM

**Title:** Discrimination of the babbling of infants from different language environments

**Principal Investigator:** Daisy Edmondson, Ph.D. Program in Psychology (Developmental), City University Graduate School

**Purpose:** The purpose of this study is to determine if participants can discriminate between the babbling of infants from different language environments.

**Procedure:** You will be asked to listen to a tape recording of infants babbling. The infants on the tape will come from homes in which different languages are spoken by the parents. You will be asked to identify which samples come from the English-exposed infants and which do not. You may listen to the tape as many times as you wish.

**Rights of Participants:** Participation is entirely voluntary, and you may withdraw your consent and have the testing session discontinued at any time. Access to the data collected will be restricted to authorized research staff. You will not be identified by name in any future presentations or publications of this data. If you have any questions regarding this research, please call Daisy Edmondson at 212-642-2556. If you have any questions concerning your rights as a participant in this study, you can call the Office of Sponsored Research, City University of New York, at 212-642-2059.

I agree to participate in the study described above:

Signature \_\_\_\_\_ Date \_\_\_\_\_

Researcher's Signature \_\_\_\_\_ Date \_\_\_\_\_

**RESPONSE SHEET**

Participant Number: \_\_\_\_\_ Phonetic Training (Y/N): \_\_\_\_\_

Children/ages: \_\_\_\_\_ Tape version: \_\_\_\_\_

Language background (first language other than English; other languages spoken/studied--how fluent/how many years of study):  
\_\_\_\_\_  
\_\_\_\_\_

TRIAL 1	
TRIAL 2	
TRIAL 3	
TRIAL 4	
TRIAL 5	
TRIAL 6	
TRIAL 7	
TRIAL 8	

**RESPONSE SHEET**

Participant number: \_\_\_\_\_

Date: \_\_\_\_\_

Participant's age: \_\_\_\_\_

Sex: \_\_\_\_\_

Phonetic training: \_\_\_\_\_

Musical training: \_\_\_\_\_

Children: \_\_\_\_\_

Race/ethnicity: \_\_\_\_\_

Language  
background/experience: \_\_\_\_\_

trial 1	
trial 2	
trial 3	
trial 4	
trial 5	
trial 6	
trial 7	
trial 8	

## APPENDIX B

### Differences in the Number of Consonants per Sample Set

Five different sets of babbling samples were used in these studies. The mean number of consonants varied as a function of sample set i.e., the mean number of consonants in the sample set designated as "English 1" was 10.75, while the mean number of consonants in the French sample set was 4.25 (see Table 20).

To see if the differences in the means of the five sample sets were statistically significant, a one-way ANOVA was performed. The counts shown in Table 20 were transformed as follows, prior to being used in the ANOVA: First, the logs of each count were taken and 1 was added so that none of the counts would equal zero; the resulting figures were then multiplied by 10 to make the numbers larger and more manageable.

The transformed data are shown below in Table 21. The ANOVA yielded significant results:  $F(4,15) = 12.49, p < .01$ . This indicates that the differences in consonant number by language group were significant.

Protected *t*-tests were then performed on all language pairs that were tested in Experiments 1 through 8. The results show that significant differences in number of consonants existed between English 1 and Arabic, Chinese, French, and English 2; the differences between English 2 and Arabic and between English 2 and Chinese were not significant (see Table 22).

Table 20

Consonant Count per Speech Sample, by Language Group

	Eng1	Eng2	Arabic	Chinese	French
<b>sample 1</b>	6	1	0	5	6
<b>sample 2</b>	8	1	0	2	5
<b>sample 3</b>	15	0	0	0	0
<b>sample 4</b>	12	1	0	1	6
<b>mean</b>	10.25	0.75	0	2	4.25

Table 21

Log-Transformed Consonant Counts, Means, and Standard Deviations

	<b>Eng1</b>	<b>Eng2</b>	<b>Arabic</b>	<b>Chinese</b>	<b>French</b>
<b>sample 1</b>	1.78	1	1	1.69	1.78
<b>sample 2</b>	1.9	1	1	1.3	1.69
<b>sample 3</b>	2.18	1	1	1	1
<b>sample 4</b>	2.08	1	1	1	1.78
<b>mean</b>	1.98	1	1	1.25	1.56
<b>s.d.</b>	0.18	0	0	0.33	0.38

Table 22

Differences Between Sets of Language Samples Tested in Experiments 1-8

<b>Language Pair</b>	<b>Protected <i>t</i> Value (d.f.= 7)</b>	<b>Two-Tailed Probability</b>
Eng1-Arabic	5.86	<.001
Eng1-Chinese	4.39	<.005
Eng1-French	2.51	<.05
Eng1-Eng2	5.86	<.001
Eng2-Arabic	0	n.s.
Eng2-Chinese	1.48	n.s.

**REFERENCES**

- Aslin, R. N. (1987). Visual and auditory development in infancy. In J. D. Osofsky (Ed.), *Handbook of infant development* (2nd ed., pp. 5-97). New York: Wiley.
- Atkinson, K., MacWhinney, B., & Stoel, C. (1968). *An experiment on the recognition of babbling* (Report No. WP-15). Berkeley, CA: Language-Behavior Research Laboratory.
- Bel'tyukov, V. I., & Salakhova, A. D. (1973). Babbling in the hearing child. *Voprosy Psikhologii*, 2, 587-605.
- Boysson-Bardies, B. de, Bacri, N., Sagart, L., Poizat, M. (1981). Timing in late babbling. *Journal of Child Language*, 8, 525-539.
- Boysson-Bardies, B. de, Hallé, P., Sagart, L., & Durand, C. (1989). A crosslinguistic investigation of vowel formants in babbling. *Journal of Child Language*, 16, 1-17.
- Boysson-Bardies, B. de, Sagart, L., Bacri, N. (1981). Phonetic analysis of late babbling: A case study of a French child. *Journal of Child Language*, 8, 511-524.
- Boysson-Bardies, B. de, Sagart, L., & Durand, C. (1984). Discernible differences in the babbling of infants according to target language. *Journal of Child Language*, 11, 1-15.
- Boysson-Bardies, B. de, & Vihman, M. M. (1991). Adaptation to language: Evidence from babbling and first words in four languages. *Language*, 67(2), 297-319.

Brown, R. (1958). *Words and things*. Glencoe, IL: Free Press.

Cooper, R. P., & Aslin, R. N. (1989). The language environment of the young infant: Implications for early perceptual development. *Canadian Journal of Psychology*, *43*, 247-265.

Crook, C. (1978). Taste perception in the newborn infant. *Infant Behavior and Development*, *1*, 52-59

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, *208*, 174-176.

DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development*, *9*, 133-150.

Delattre, P. C. (1961). La leçon d'intonation de Simone de Beauvoir, étude d'intonation déclarative comparée. *The French Review*, *35*, 59-67.

Hallé, P. A., Boysson-Bardies, B. de., & Vihman, M. M. (1991). Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by Japanese and French infant. *Language and Speech*, *34*(4), 299-318.

Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Wright Cassidy, K., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, *26*, 269-286.

Holmgren, K., Lindblom, B., Aurelius, G. Jalling, B., & Zetterström, R. (1986). On the phonetics of infant vocalization. In B. Lindblom & R. Zetterström (Eds.),

*Precursors of early speech* (pp. 51-63). New York: Stockton Press.

Hooker, D. (1969). *The prenatal origins of behavior*. New York: Hafner.

Jakobson, R. (1968). *Child language, aphasia, and phonological universals* (A.R Keiler, Trans.). The Hague: Mouton. (Original work published 1941)

James, W. (1890). *The principles of psychology*. New York: Dover Publications.

Jusczyk, P. W. (1989, April). *Perception of cues to clausal units in native and non-native languages*. Paper presented at the meeting of the Society for Research in Child Development, Kansas City, MO.

Jusczyk, P. W. (1992). Developing phonological categories from the speech signal. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 17-64). Timonium, MD: York Press.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development, 64*, 675-687.

Jusczyk, P. W., Friederici, A. D., Wessels, J. M. I., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language, 32*, 402-420.

Jusczyk, P. W., Hirsh-Pasek, K., Kemler Nelson, D. G., Kennedy, L. J., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology, 24*, 252-293.

Jusczyk, P. W., Kemler Nelson, D. G., Hirsh-Pasek, K., & Schomberg, T. (1994). *Perception of acoustic correlates to clausal units in a foreign language by American infants*. Manuscript in preparation.

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

Kemler Nelson, D. G. (1989, April). *Developmental trends in infants' sensitivity to prosodic cues correlated with linguistic units*. Paper presented at the meeting of the Society for Research in Child Development, Kansas City, MO.

Kemler Nelson, D. G., Hirsh-Pasek, K., Jusczyk, P. W., & Wright Cassidy, K. (1989). How the prosodic cues in motherese might assist in language learning. *Journal of Child Language*, *16*, 53-68.

Kent, R. D. (1992). The biology of phonological development. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 65-90). Timonium, MD: York Press.

Kent, R. D., & Read, C. (1992). *The acoustic analysis of speech*. San Diego: Singular Publishing Group.

Koopmans-van Beinum, F. J., & van der Stelt, J. M. (1986). Early stages in the development of speech movements. In B. Lindblom & R. Zetterström (Eds.), *Precursors of early speech* (pp. 37-50). New York: Stockton Press.

Kuehn, G. B. A. (1985). Discrimination and identification of Chinese and

English infant vocalizations (Doctoral dissertation, Washington University, 1985).

*Dissertation Abstracts International*, ?, xxxx.

Levitt, A. G., & Wang, Q. (1991). Evidence for language-specific rhythmic influences in the reduplicative babbling of French- and English learning infants.

*Language and Speech*, 34(3), 235-249.

Lieberman, P., Crelin, E. S., & Klatt, D.H. (1972). Phonetic ability and related anatomy of the newborn, adult human, Neanderthal man, and the chimpanzee.

*American Anthropologist*, 74, 287-307.

Locke, J. L. (1983). *Phonological acquisition and change*. New York: Academic Press.

Locke, J. L. (1989). Babbling and early speech: Continuity and individual differences. *First Language*, 9, 191-206.

Locke, J. L., & Pearson, D. M. (1992). Vocal learning and the emergence of phonological capacity: A neurobiological approach. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 91-130). Timonium, MD: York Press.

MacFarlane, A. (1978). Olfaction in the development of social preferences in the human neonate. In R. Porter & M. O'Connor (Eds.), *Parent-infant interaction* (pp. 103-117). Amsterdam: Elsevier/Excerpta Medica/North-Holland.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29,

143-178.

Mitchell, P. R., & Kent, R. D. (1990). Phonetic variation in multisyllabic babbling. *Journal of Child Language*, *17*, 247-265.

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, *16*(4), 495-500.

Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology, Vol. 1: Production* (pp. 93-112). New York: Academic Press.

Oller, D. K., & Eilers, R. E. (1982). Similarity of babbling in Spanish- and English-learning babies. *Journal of Child Language*, *9*, 565-567.

Olney, R. L., & Scholnick, E. K. (1976). Adult judgments of age and linguistic differences in infant vocalizations. *Journal of Child Language*, *3*, 145-155.

Petinou, K., Schwartz, R. G., Mody, M., & Gravel, J. S. (in press). The impact of otitis media with effusion on early phonetic inventories: A longitudinal prospective investigation. *Clinical Linguistics & Phonetics*.

Polka, L., & Werker, J. F. (in press). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*.

Querleu, D., Renard, X., Versyp, F., Paris-Delrue, L., & Crèpin, G. (1988). Fetal hearing. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, *29*, 191-212.

Ramsay, D. (1984). Onset of duplicated syllable babbling and unimanual handedness in infancy: Evidence for developmental change in hemispheric specialization? *Developmental Psychology*, *20*, 64-71.

Roug, L., Landberg, I., & Lundberg, L.-J. (1989). Phonetic development in early infancy: A study of four Swedish children during the first eighteen months of life. *Journal of Child Language*, *16*, 19-40.

Sasaki, C. T., Levine, P.A., Laitman, J. T., & Crelin, E. S. (1977). Postnatal descent of the epiglottis in man. *Archives of Otolaryngology*, *103*, 169-171.

Stark, R. E. (1980). Stages of speech development in the first year of life. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology, Vol. 1: Production* (pp. 73-92). New York: Academic Press.

Steiner, J. (1979). Human facial expressions in response to taste and smell stimulation. In H. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behavior: Vol.13* (pp. 257-295). New York: Academic Press.

Thelen, E. (1981). Rhythmical behavior in infancy: An ethological perspective. *Developmental Psychology*, *17*, 237-257.

Thelen, E. (1991). Motor aspects of emergent speech: A dynamic approach. In N. A. Krasnegor, D. M. Rumbaugh, R. L. Schiefelbusch, & M. Studdert-Kennedy (Eds.), *Biological and behavioral determinants of language development* (pp. 339-362). Hillsdale, NJ: Lawrence Erlbaum.

Turkewitz, G. (1988). A prenatal source for the development of hemispheric

specialization. In D. L. Molfese & S. J. Segalowitz (Eds.), *Developmental implications of brain lateralization* (pp. 73-81). New York: Guilford Press.

Vihman, M. M. (1993). Variable paths to early word production. *Journal of Phonetics*, 21, 61-82.

Vihman, M. M., & Boysson-Bardies, B. de (1994). The nature and origins of ambient language influence on infant vocal production and early words. *Phonetica*, 51, 159-169.

Vihman, M. M., DePaolis, R. A., & Davis, B. L. (1998). Is there a "trochaic bias" in early word learning? Evidence from infant production in English and French. *Child Development*, 69(4), 935-949.

Vihman, M. M., Macken, M. A., Miller, R., Simmons, H., & Miller, J. (1985). From babbling to speech: A reassessment of the continuity issue. *Language*, 61, 395-443.

Vihman, M. M., & Miller, R. (1988). Words and babble at the threshold of language acquisition. In M. D. Smith & J. L. Locke (Eds.), *The emergent lexicon* (pp. 151-184). New York: Academic Press.

WaveStudio 2.0 [Computer software]. (1994). Milpitas, CA: Creative Labs.

Weir, R. (1966). Some questions on the child's learning of phonology. In F. Smith & G. A. Miller (Eds.), *The genesis of language: a psycholinguistic approach* (pp. 153-168). Cambridge, MA: M.I.T. Press.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception:

Initial capabilities and developmental change. *Developmental Psychology*, 24(5), 672-683.

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

Whalen, D. H., Levitt, A. G., & Wang, Q. (1991). Intonational differences between reduplicative babbling of French- and English-learning infants. *Journal of Child Language*, 18, 501-516.