

**Developmental Changes in Arabic Babbling in Relation to English and French Babbling**

Abdulsalam Alhaidary

School of Communication Sciences and Disorders

McGill University, Montreal, Canada

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree  
of Doctor of Philosophy.

November 2012

© Abdulsalam Alhaidary, 2012

## Table of contents

ABSTRACT.....	i
Résumé.....	iii
Statement of Originality.....	vi
Acknowledgements.....	vii
1. INTRODUCTION .....	1
1.1 An overview.....	1
1.2 Literature review .....	4
1.2.1 Development of speech perception during infancy .....	4
1.2.2 Development of speech production during infancy .....	18
1.2.3 Variables influencing infant speech production .....	25
1.2.3.1 Biological influences .....	26
1.2.3.2 Linguistic environmental influences.....	34
1.3 Purpose of the study.....	57
2. ARABIC .....	60
2.1 Background on the Arabic language.....	60
2.2 Arabic phonology.....	62
3. METHOD .....	69
3.1 Participants.....	69
3.2 Apparatus .....	71
3.3 Procedure .....	71
3.4 Method of Study One.....	76
3.5 Method of Study Two .....	85
4. RESULTS .....	88
4.1 Results of Study One .....	88
4.1.1 Reliability.....	88
4.1.2 The analyses of the center and corners of the vowel space: Arabic versus French .....	89
4.1.3 The analyses of the center and corners of the vowel space: Arabic versus English..	101
4.1.4 The analyses of the size of the vowel space: Arabic versus French versus English..	112
4.2 Results of Study Two.....	116
4.2.1 Reliability.....	116

4.2.2 Results for place of articulation .....	116
4.2.3 Results for manner of articulation.....	121
5. DISCUSSION .....	125
5.1. An overview of the research objectives .....	125
5.2. A summary and interpretations of the findings.....	127
5.3 Implications.....	148
5.4 Limitations .....	150
5.5 Future directions .....	152
5.6 Conclusion .....	154
REFERENCES .....	156

## ABSTRACT

Infant vocalization undergoes dramatic changes during the transition from newborn cooing and infant babbling toward the production of meaningful speech. However, the processes underlying these changes are not fully understood, particularly at the age at which ambient language input begins to influence infant babbling. The aim of this study was to describe the babbling produced by Arabic infants—using acoustic and phonetic metrics—and then examine the effects of linguistic environmental input on babbling by comparing the babble produced by Arabic infants to that produced by English and French infants. Speech samples were collected from infants learning Arabic ( $N = 31$ ; age: 281-591 days), English ( $N = 20$ ; age: 303-553 days) or French ( $N = 23$ ; age: 311-566 days). Each utterance was transcribed according to the International Phonetic Alphabet and then coded according to infraphonological categories.

Two studies were conducted. Study one aimed to describe the vowel space of Arabic infants and then compare it to the vowel space of English and French infants. First (F1) and second (F2) formant frequencies were identified in all vowels produced with full resonance, normal phonation, and speech-like duration. These F1 and F2 frequencies were used to calculate the compact-diffuse ( $F2 - F1$ ) and grave-acute ( $(F2+F1)/2$ ) values for each vowel so that the extreme corners of each infant's vowel space could be identified. Multiple regression and analysis of variance analyses were used to examine the effects of language and age on the infant vowel space (centre, corners, and area of the space). Developmental changes in the expansion of the vowel space toward the grave corner were observed in all language groups. In addition, a language-specific pattern of changes on the

vowel space were observed with age: Arabic infants showed unchanged F1 and F2 values at the centre of the space, expansion toward the compact corner, a larger vowel space than infants in the other language groups at all ages studied; French infants showed a decreased F1 and unchanged F2 values at the centre, an expansion toward the diffuse corner, and a shrinking from the acute corner toward the center of the vowel space; and English infants showed decreased F2 values but unchanged F1 values at the centre of the vowel space. Study two aimed to describe the consonantal repertoires of Arabic infants and then compare these consonantal repertoires across three infant language groups. Infants were organized into three age groups (10-12, 12-15 and 15-18 months), and the consonants produced in canonical (CS) and marginal (MS) syllables with normal phonation were grouped into manner and place categories. Analysis of variance revealed no significant crosslinguistic differences in the frequency of production of any manner or place categories.

This study showed developmental and language-specific changes in the infant vowel space when vowels were submitted to acoustic analysis. However, early crosslinguistic differences in vowel space were not accompanied by crosslinguistic differences in the consonantal repertoires based on phonetic transcription analyses. The overall findings from the current study provide evidence for the *interactional hypothesis* and suggest that the development of infant babbling is influenced by a complex interaction of endogenous and exogenous processes including biological development of the vocal tract and language input from the ambient environment.

## Résumé

La vocalisation du nourrisson subit des changements importants, du gazouillis du nouveau-né et du babillage du bébé vers la production de la parole significative. Cependant, on ne comprend pas entièrement les processus sous-jacents à ces changements, en particulier l'âge auquel l'acquisition de la langue ambiante commence à influencer le babillage. L'objectif principal de cette étude était de décrire le babillage des tout-petits apprenant l'arabe à l'aide de mesures acoustiques et phonétiques et d'examiner les effets de l'apport de l'environnement linguistique sur le babillage infantile en comparant le babil produit par les jeunes enfants arabes à celui d'enfants apprenant l'anglais ou le français. Des échantillons de parole ont été recueillis auprès de jeunes enfants apprenant l'arabe (N=31; tranche d'âge: 281-591 jours), l'anglais (N=20; tranche d'âge: 303-553 jours) ou le français N=23; tranche d'âge: 311 à 566 jours). Chaque énoncé a été transcrit conformément à l'API, puis codé selon les catégories infraphonologiques.

Deux expériences ont été menées. La première visait à décrire l'espace vocalique acoustique des jeunes enfants arabes, et de le comparer à celui de jeunes enfants anglais et français. Les 1<sup>res</sup> (F1) et 2<sup>es</sup> (F2) fréquences de formants ont été identifiées dans toutes les voyelles produites avec pleine résonance, phonation normale et durée comme celle de la parole. F1 et F2 ont été utilisées pour calculer les valeurs compact-diffus ( $F2 - F1$ ) et grave-aigu ( $[F2 + F1] / 2$ ) pour chaque voyelle, pour identifier les coins extrêmes de l'espace vocalique de chaque enfant. La régression multiple et l'observation des analyses de variance ont été utilisées pour examiner les effets de la langue et de l'âge sur l'espace vocalique infantile (centre, coins, et zone de l'espace). Des changements

développementaux dans l'expansion de l'espace vocalique vers le coin grave ont été observés dans tous les groupes linguistiques. De plus, un modèle de changements spécifique au langage de l'espace vocalique a été observé avec l'âge: les enfants arabes ont montré des valeurs F1 et F2 inchangées au centre de l'espace, une expansion vers le coin compact, un espace vocalique plus grand que les enfants des autres groupes de tous les âges étudiés. Les enfants français ont montré une diminution de F1 et des valeurs F2 inchangées au centre, une expansion vers le coin diffus, et un rétrécissement du coin aigu vers le centre de l'espace vocalique et les enfants anglais ont montré une diminution des valeurs F2 et des valeurs F1 inchangées au centre de l'espace vocalique. La deuxième expérience visait à décrire les répertoires consonantiques des enfants arabes et de comparer ces répertoires consonantiques dans trois groupes linguistiques de jeunes enfants. Les consonnes qui ont été produites dans les syllabes canoniques (CS) et marginales (MS) avec la phonation normale ont été incluses dans l'analyse. Les enfants ont été divisés en trois groupes (10-12, 12-15 et 15-18 mois), et les consonnes, regroupées en catégories phonétiques en fonction de la manière et du lieu de production. L'analyse de variance n'a pas révélé de différences significatives translinguistiques dans la fréquence de production pour toutes catégories de manière ou lieu.

L'étude a montré des changements développementaux et spécifiques au langage dans l'espace vocalique des tout-petits quand les voyelles ont été soumises à une analyse acoustique. Cependant, des différences translinguistiques précoces dans l'espace vocalique n'étaient pas accompagnées par des différences translinguistiques dans les répertoires consonantiques selon les analyses de transcription phonétique. L'étude actuelle a fourni des preuves de l'hypothèse interactionnelle et suggère que le développement du babillage

infantile est influencé par une interaction complexe des processus endogènes et exogènes, notamment le développement biologique du conduit vocal et la langue de l'environnement ambiant.



## **Statement of Originality**

This dissertation is an original scholarly investigation that traces the speech development of Arabic infants from 10 to 18 months of age. The investigation provided original developmental data on Arabic babbling: there are no other published phonetic studies that describe prelinguistic vocalizations produced by Arabic-learning infants of this age range using a detailed acoustic analysis. Furthermore, the sample of 31 infants described in this report is unusually large. The analyses of these data contribute to our knowledge about the course of Arabic speech development. The crosslinguistic comparisons of monolingual Arabic babbling to the babbling of monolingual English infants and monolingual French infants yielded original evidence of linguistic environmental influences on babbling. Furthermore, the results confirmed the hypothesis that developmental expansion of infant vowel space would be related to the complexity of the vowel system of ambient language. Specifically, Arabic infants achieved an early expansion of the vowel space compared to the English and French infants as indicated by the larger size of the triangular vowel space and more extreme values in diffuseness, graveness, and compactness features from the age of 10 months. The findings contribute to current understanding of the mechanisms that underlie early speech development by highlighting the role of the ambient language environment generally and the complexity of the input vowel space specifically. Some of the results of this study were presented at the November 2010 conference of the American Speech-Language and Hearing Association (Philadelphia, PA) and at the April 2011 Research in Child Development Biennial Meeting (Montreal, QC).

## Acknowledgements

First and foremost, I thank Allah for his blessings for me to complete my study and finish this project. I would like to express my deep gratitude to my supervisor Dr. Susan Rvachew for teaching and training me, for sharing her understanding and expertise with me, for her unlimited support throughout my study, and for being such a great teacher in all that entails. Also, I thank my thesis committee Dr. Linda Polka and Dr. Elin Thordardottir for their feedback and comments on earlier drafts. I also thank the administrative Staff in Beatty Hall (Antoinette, Lili, Miriam, and Karen) with a special thanks to Ms. Antoinette Sommer for her ongoing support, kindness and friendliness. I would like to thank Marie Desmarteau and Kanwar Anit Singh Saini for their technical help with the speech sample recordings, Bader Alqabani for his help with the hearing screening equipment in Saudi Arabia, Wayne Egers for proofreading help, Franck Bélanger for the French translation of the abstract and Young Ja Nam for the reliability check. I also would like to thank my fellow PhD students Françoise Brosseau-Lapr  and Kathrin Rees for their friendship. I would like to acknowledge the support from my family: thanks to my parents for their unlimited support and prayers and special thanks to my wife Alaa for her continuous support and encouragement throughout my study. I also would like to express my gratitude to the parents and babies who welcomed me at their homes and agreed to participate in this study. In addition, I thank the Saudi Arabian Cultural Bureau in Ottawa for supervising and facilitating my study in Canada. This research was supported by grants from Natural Sciences and Engineering Research Council of Canada and the Canadian Language and Literacy Research Network. Last, I would like to acknowledge the scholarship from the government of Saudi Arabia through King Saud University that supported me in completing my study.

# **1. INTRODUCTION**

## **1.1 An overview**

Much of the past research on infant speech development has been conducted in English. Thus, data from infants learning languages other than English are needed to confirm and extend the earlier findings about the different issues associated with speech development. For example, one issue that needs further exploration is the role of linguistic environmental influences on infant early speech production. No doubt exists that language-specific exposure influences the vocal patterns of infants as they mature; however, the age at which babbling departs from universal patterns and begins to be influenced by the ambient language is debatable.

Over the past 30 years, research on early vocal development has been influenced by two prominent but conflicting views on the development of babbling. One line of research focuses on the similarity of babbling across different linguistic communities, following from Jakobson's (1941/1968) view of the universality of infant vocalizations: Jakobson argued that infants babble similarly (regardless of language-specific exposure) and produce a great variety of sounds, some of which are not produced in the background language and may occur in other languages to which the infants are not exposed. A phonetic explanation (i.e., biological constraints arising from an immature articulatory system during infancy) is proposed to justify the similarity in crosslinguistic babbling (e.g., MacNeilage & Davis, 2000, 2001). In contrast, another line of research has argued that infant vocal behaviour is influenced by language input and has searched for evidence of early crosslinguistic differences in babbling. For example, de Boysson-Bardies, Halle,

Sagart, and Durand (1989) proposed that an interaction between ambient language and biological influences accounts for the changes in infant babbling (*interactional hypothesis*). As will be related in the literature review to follow, researchers working from this perspective have attempted to show that infants gradually replicate unique aspects of the ambient language input. The idea that babbling and adult speech input is similar is consistent with the *babbling drift hypothesis*. Brown (1958) was one of the first scholars to discuss “the fact that... [babbling] drifts in the direction of the speech the infant hears” (p. 199). However, I will argue that the developmental process of infant speech production is more complex than the assumption of the *babbling drift hypothesis*, that is, that babbling simply drifts toward the adult language model. As suggested by the *interactional hypothesis* (de Boysson-Bardies et al., 1989), infant speech production is explained by an interaction among many variables—among them ambient language input and the maturation of the articulatory structures and functions. The active role of the infant language learner must be given greater prominence in this interactional account with infant attentional preferences as shaped by the complexity of the input also playing a role.

The present study investigates developmental changes in the babbling of monolingual infants learning Arabic, and compares these changes to monolingual infants learning English or French. These three languages have substantial differences in their phonological and phonetic systems. Particularly, the complexity of the adult vowel systems differs across the three languages (i.e., Arabic has a simple vowel inventory compared to the more complex vowel inventories of English and French). This crosslinguistic comparison permits an investigation of the early effects of ambient

language input on the size and expansion of the vowel space. Previously, research has demonstrated that the expansion of the vowel space is universal in normal hearing infants, regardless of the language they are learning (Kent & Murray, 1982; Rvachew, Mattock, Polka, & Menard, 2006). In this study, it is expected that there will be language-specific influences on this process however. Specifically, a less complex vowel space in the input language should facilitate the achievement of a larger vowel space at a younger age; on the other hand, a more complex vowel space in the input language should lead to a comparatively smaller vowel space due to lesser perceptual salience of the corner vowels for infants learning these languages. To test this hypothesis, the current study used an investigative approach that included the search for a specific ambient language effect (crosslinguistic differences in vowel space area) rather than a drift toward the ambient language or a match between babbling and adult speech. Furthermore, the research design involved a large sample of infants, an infraphonological description of infant vocalizations, phonetic transcription, and a detailed acoustic description of the infant vowel space. The use of such a method may help to overcome some of the issues identified in the earlier research on ambient language influences on babbling.

The remainder of this chapter describes the relevant literature on the influences of the linguistic environment on infant babbling. Section 1.2.1 describes the development of infant speech perception, including auditory mechanisms. Section 1.2.2 reviews the literature on the development of speech production during infancy. Section 1.2.3 discusses the biological and environmental variables associated with infant babbling. Section 1.3 explains the purpose of this study and provides the research questions.

## **1.2 Literature review**

The research on the influences of ambient language input on the infant's path to acquiring household language requires an understanding of the infant's remarkable speech perception skills, infant speech production development, infant articulatory system maturation, and the role of speech input on infant vocalizations. In addition to other variables (such as social interaction and cognitive development), these variables just mentioned are interrelated in a dynamic process throughout the course of infant speech development. This literature review brings together a body of research on these aspects of infant speech development, which highlights the significant contribution of the present investigation.

### **1.2.1 Development of speech perception during infancy**

The purpose of this section is to review the literature on the development of speech perception during infancy. It begins with a discussion about the anatomical immaturity of an infant's auditory system and its influences on some auditory functions important for processing auditory input. Next, the development of speech perception, and the role of language-specific exposure during infancy is discussed.

#### *Anatomical and functional maturities of the auditory mechanism*

The capacity to hear and perceive sounds depends on the maturation of the auditory mechanism. The different structures of an infant's auditory mechanism reach maturation at different timelines that range from prenatal to late childhood. One of the earliest auditory structures to achieve maturation in terms of its size and shape is the

cochlea. Bredberg (1967) conducted an autopsy investigation on the human cochlea from infancy to old age and found that at two and a half months of fetal age, the sensory cells (inner and outer hair cells) in the cochlea are not separated from the supporting cells in the organ of Corti. The cochlea starts to coil at the basal end and advances gradually to the apex end. From three to six months of fetal age, the cochlea enters a rapid pace of growth in which the sensory hearing cells evolve and the organ of Corti forms. By the sixth fetal month, the cochlea is observed to be adult-like in terms of shape, size, and cellular patterns (e.g., the number of rows of inner and outer hair cells, and hair cells connection to spiral ganglion). Subsequent research (e.g., Bibas et al., 2008; Lavigne-Rebillard & Pujol, 1987; J. K. Moore & Linthicum, 2007; Pujol & Lavigne-Rebillard, 1992) confirmed this pattern of cochlea maturation. However, Bibas et al. (2008) reported that some cochlear structures are not fully developed at the sixth fetal month (i.e., 25 weeks of gestational age). According to Bibas et al. (2008), the tunnel of Corti continues to develop until near birth. Other cochlea structures also continue to develop throughout the gestation period, for example, supporting cells increase in size and become columnar-like in shape. Overall, the human cochlea and most of its structures seem to be adult-like in terms of shape, size, and innervation by the end of the second trimester of pregnancy.

Unlike the prenatal maturation of the cochlea, other auditory structures reach maturation at a later age. Moore and her associates (e.g., J. K. Moore, 2002; J. K. Moore & Linthicum, 2007; J. K. Moore, Perazzo, & Braun, 1995) conducted a series of investigations and literature reviews on the course of the maturation of various cortical and subcortical auditory structures. They found that these structures continue to develop

after birth to late childhood. For example, myelination was first observed around 26 weeks of gestational age, and the density of the myelin sheath continued to increase, reaching an adult-like appearance by 12 months of age (J. K. Moore, 2002; J. K. Moore & Linthicum, 2007). Axons in the auditory cortex also were apparent only at the superficial layers of the auditory cortex during the perinatal period, and they continue to extend horizontally over the auditory cortex and project vertically to the deeper cortical layers as age increases, reaching maturation in terms of axonal density and cortical and subcortical connection at 12 years of age (J. K. Moore et al., 1995). Thus, the auditory mechanism is ready to perceive speech sounds during the third trimester of pregnancy, since the organ of hearing (cochlea) is structurally developed and capable of converting the disturbances of the basal membrane into nerve signals that then transmit through the spiral ganglion and auditory pathways to the auditory cortex. However, the maturity of many other auditory structures is achieved over a prolonged period of time that extends for some structures to late childhood.

Similar to the immaturity of the auditory structures at birth, the auditory functions that are important for speech perception also undergo developmental changes after birth. Importantly, intensity, frequency, and temporal processing are among the fundamental functions for processing complex sounds such as speech (Werner, 2002).

Several studies that have investigated infants' hearing sensitivity report that infant hearing threshold is higher than the adult threshold. Based on behavioural hearing evaluations, some scholars have estimated that the infant hearing threshold is higher (poorer) than the adult threshold by 15-20 dB SPL (at 250 Hz-4 KHz) and 30 dB SPL (at



8 KHz) for infants aged three months. At 6 months of age, the difference decreases (except for the 250 Hz range) and further decreases from 6 to 12 months of age to become within 10-15 dB SPL of the adult threshold (Olsho, Koch, Carter, Halpin, & Spetner, 1988; Tharpe & Ashmead, 2001). These findings are comparable with other electrophysiological studies of hearing sensitivity. For example, Rance and Tomlin (2006) showed a significant decline in the mean threshold of almost 10 dB HL from birth to six weeks of age (e.g., 52.4 dB HL to 41.3 dB HL at 500 Hz) as measured by recording auditory steady-state responses (evoked neural responses to steady stimuli). Nevertheless, all the reports on the infant hearing threshold across studies are mainly estimates of the threshold (not an accurate measure) because the infants studied are too young to participate in regular audiological tests. Despite this limitation and the methodological differences across studies, agreement occurs throughout the literature that the infant hearing threshold is not adult-like and continues to develop during infancy.

Similar to auditory sensitivity, frequency processing continues to improve during infancy. Two functions are commonly investigated to examine this processing: frequency resolution and frequency discrimination. *Frequency resolution* (i.e., the ability to detect or resolve one tone from other competing tones based on frequency [Saffran, Werker, & Werner, 2006]) is not mature until six months of age. This function can be tested by measuring the listener's ability to detect a stimulus while background noise is present (i.e., the masking threshold). A number of studies have investigated the development of frequency resolution. For example, the investigations of the *tuning curves* of the auditory system (i.e., a graphical representation of how the basilar membrane responds to a target tone when presented with simultaneous tones or noises)—as measured by

electrophysiological measures (Abdala & Folsom, 1995) and behavioural measures (Spetner & Olsho, 1990)—showed that frequency resolution is not mature in neonates. Three-month-old infants have a broader and less steep tuning curve for high frequencies, whereas six-month-old infants have a curve pattern similar to an adult tuning curve for high and low frequencies (the broader the curve is, the less tuned the system for a particular tone). Similar to frequency resolution, the immaturity of frequency discrimination also has been reported. Novitski and her associates (Novitski, Huotilainen, Tervaniemi, Näätänen, & Fellman, 2007; Novitski, Tervaniemi, Huotilainen, & Näätänen, 2004) examined the 250-4000 Hz frequency discrimination of newborns as determined by mismatch response (MMR) and of adults as determined by mismatch negativity (MMN)—both commonly-used electrophysiological measures of brain responses to changes in sound. The polarity of the mismatch response in infants may be negative or positive, hence MMR. The experiment was conducted by using an oddball paradigm in which standard (repetitive) sounds (e.g., a 250 Hz tone) were interrupted by deviant (infrequent) sounds with a  $\pm 5\%$  and  $20\%$  change in the original tone. The infants' MMR was recorded for the  $\pm 20\%$  changes, and the adults' MMN was recorded for the  $\pm 5\%$  changes, suggesting an immaturity of frequency discrimination during infancy. Thus, the frequency coding of the basilar membrane is not adult-like in part because of the infants' limited experiences with sounds prenatally and during the first months of life.

*Temporal processing*—the ability to process frequency and intensity changes over time—is also not adult-like during the first year of life. Researchers measure this function by examining the duration of an acoustic gap or interruption that someone can detect. The

shorter the gap is, the better the temporal processing. Adults are able to detect an acoustic gap with a duration of 5 ms depending on the frequency, whereas infants are found to have a higher gap detection threshold (Smith, Trainor, & Shore, 2006; Werner, Marean, Halpin, Spetner, & Gillenwater, 1992). This was not only true for the 6- to 7.5-month-old infants in Smith et al.'s (2006) study but also for the 12-month-old infants in Werner et al.'s (1992) study. The 12-month-old groups had a lower gap-detection threshold compared to the younger infant groups but still were not adult-like, suggesting that this age group has an improvement in temporal processing. Thus, infants have a higher gap-detection threshold compared to adults, which suggests that infant's temporal processing of acoustic signals is not adult-like.

Taken together, the findings from the just-cited studies about the anatomical and functional changes in the auditory mechanism during infancy suggest that infants do not attend to or have access to the spectral and temporal information of speech in an adult-like manner due, in part, to the constraints imposed by their immature auditory structures and functions. For example, the perception of high-frequency sounds (e.g., fricatives) may be constrained by an infant's reduced efficiency of frequency processing at high frequencies. Researchers have shown that infants find some fricative contrasts difficult to discriminate and that their perception of these contrasts may be based on dynamic cues (transition information) in the syllable rather than the spectral characteristics of the fricative noise (Eilers, Wilson, & Moore, 1977; Levitt, Jusczyk, Murray, & Carden, 1988; Nittrouer, 2001). For example, infants showed some difficulty in discriminating between /f/-/θ/ contrast (Eilers et al., 1977; Levitt et al., 1988) and /f/-/s/ contrast (Nittrouer, 2001) due to the great similarity in their spectral properties.

*Prenatal hearing and postnatal speech perception*

Despite the anatomical and functional immaturity of an infant's auditory system, ample evidence shows that infants are remarkably capable of perceiving and processing speech input, although not necessarily in an adult-like manner. Infant's experiences with speech sounds starts prenatally, particularly during the last trimester. Techniques to investigate prenatal hearing are limited to observing how fetal movements, heart rate, and breathing (movements of the diaphragm, thorax, and abdomen) change when presented with sound stimuli (Kisilevsky & Low, 1998). Kisilevsky and her colleagues conducted a series of investigations on the hearing sensitivity of low- and high-risk fetuses for premature delivery between 27 to 41 weeks of gestational age (e.g., Kisilevsky et al., 2003; Kisilevsky, Pang, & Hains, 2000). The stimuli used consisted of bursts of white noise, a two minute speech segment of their mother's voice and a two minute speech segment of an unfamiliar female from 90 to 115 dB SPL delivered through a loudspeaker placed near the mothers' abdomens. Changes to heart rate and body movement following the stimulus onset were observed at 30 weeks of gestational age for both groups of fetuses. Kisilevsky and her associates' findings are consistent with Hepper and Shahidullah (1994) who found evidence of fetal responses to pure tones of low to high frequencies from 19 to 35 weeks of gestational age. Between 19 to 27 weeks of gestational age, fetuses responded to low-frequency tones (100, 250, and 500 Hz). From 29 weeks of gestational age, they responded to high-frequency tones (1000 and 3000 Hz). By 33 weeks of gestational age, all the fetuses recruited for this study (400 fetuses) responded to all the test frequencies. Thus, the auditory system is functioning before infants are born, which is due to the early maturation of the cochlea.

Furthermore, infants have the capability to detect and discriminate sound contrasts. In their seminal work, Eimas, Siqueland, Jusczyk, and Vigorito (1971) reported that one- to four-month-old infants are able to perceive the differences between the English contrast ([pa] and [ba]). These researchers constructed two experimental conditions: two within-adult-phonemic category stimuli with 20 ms of difference in the voice onset time (VOT), which is the timing of the release of stop closure and the onset of vocal fold vibrations; and two between-adult-phonemic category stimuli with 20 ms of difference in VOT. Infants heard the continuous presentation of the first stimulus until they showed a decrease in their sucking patterns, which suggests habituation. After the second stimulus was presented, any change in their sucking pattern was taken as an indication of their perception of the differences between the two stimulus presentations. The results showed a significant shift in sucking patterns when the stimulus changed for the two between-adult-phonemic category conditions. This finding also is considered to be evidence for the categorical perception of consonants. Subsequent research—using a paradigm similar to Eimas et al.’s (1971) study—show consistent findings about infants’ early discrimination ability for other consonant contrasts (e.g., Eimas, 1974; Jusczyk, Copan, & Thompson, 1978; Streeter, 1976). Similar to their capacity to distinguish consonant contrasts, infants also can discriminate among vowel contrasts. For example, two-month-old infants are able to detect the differences—as indicated by the changes in their sucking patterns—between within- and between-category vowel stimuli: within-category /i/ or /ɪ/ and between-category /i/ vs. /ɪ/ (Swoboda, Kass, Morse, & Leavitt, 1978). Moreover, six-month-old infants are able to perceive the change from /i/ to /a/ and vice versa when these sounds are produced by different speakers (male adult, female

adult, and child) with different falling and rising pitch contours (Kuhl, 1979), suggesting that six-month-old infants are able to perceive the similarity of different tokens of the same vowels and to treat them as belonging to one phonetic category. The findings from the just-cited studies (except Streeter, 1976) showed evidence that infants are able to discriminate between native speech sound contrasts.

An infant's ability to discriminate between speech sound contrasts that do not occur in their background language is also reported to occur during early infancy. For example, two-month-old Kikuyu infants are able to signal the stimulus shifts of bilabial stops from -30 ms in VOT to zero ms in VOT, and from 10 ms in VOT to 40 ms in VOT (Streeter, 1976). These phonetic contrasts do not occur in the Kikuyu adult language, which suggests that infants can make a within-phonetic category distinction. Furthermore, findings from Werker and her colleagues' (Polka & Werker, 1994; Werker & Tees, 2002) series of investigations on infant speech perception showed evidence of the infants' ability to discriminate non-native contrasts. For example, 6- to 8-month-old English infants can discriminate both English consonant contrast (/ba/-/da/), Thompson (non-English) contrast (/`ki/-/`qi/) and German vowel contrasts (tense /u/-/y/ and lax /U/-/Y/). Therefore, an infant's ability to discriminate native and non-native contrast is considered to be evidence for a universal language perception.

Furthermore, as infants grow older, their speech perception undergoes developmental changes during the last six months of their first year of life—their perception becomes language-specific, and it begins to be influenced by the ambient language input. With respect to the perception of vowels, Kuhl et al. (1992) showed that

6-month-old English and Swedish infants perceive the prototype of a native vowel category and its variants (acoustically similar but not identical) as belonging to one category; whereas they perceive the prototype of a non-native vowel category and its variants as belonging to two categories. In other words, in most trials, English infants perceive the English /i/ and its variants as the same, whereas Swedish infants tend to perceive the English /i/ and its variants as different (vice versa for the Swedish /y/ and its variants). In addition, Polka and Werker (1994) arrived at similar findings about the capacity of infants to discriminate non-native vowel contrasts. Only a few 10-12 month-old English infants reached the criteria for discriminating German contrasts (/U/ -/Y/ and /u/-/y/) compared to 6-8 month-old English infants. With respect to the perception of consonants, Werker and Tees (2002) showed that a few 8- to 10-month-old and most 10- to 12-month-old English infants in their study performed poorly in their discrimination of non-native English contrast [(/`ki/-/`qi/) and (/ʔa/-/ta/)]. Together, the just-cited studies and many others (See Houston, 2011, for a review) suggest the developmental changes in infant speech perception toward specific language perception.

However, evidence exists that older infants and adults are still sensitive to some non-native contrasts. For example, Polka, Colantonio, and Sundara (2001) compared French and English 6- to 12-month-old infants to French and English adults with respect to their ability to discriminate English consonants contrasts (/d/- /ð/) produced as an initial sound in a monosyllable word. Their findings showed that French infants and adults are able to discriminate the English contrast. However, while many French adults still reached the criterion for discrimination of the contrast (i.e., 8 correct responses out of

10 trials), French adults as a group performed less well compared to English adults. According to Polka et al., this pattern of findings suggested that specific language experience facilitates (improves) the discrimination of native contrasts and at the same time maintains the ability to discriminate non-native contrasts. Polka et al.'s (2001) findings of discriminable non-native contrasts are consistent with her earlier findings of English adults performing significantly better in discriminating German vowel contrasts (/U/-/Y/ and /u/-/y/) compared to infants from 6-12 months of age (Polka & Werker, 1994). Further, Polka et al.'s (2001) findings about the improvement in discriminating native contrasts are consistent with recent reports on developmental changes in native contrasts discrimination. For example, 4-year-old children performed better than 10- to 12-month-old infants but lower than adults in discriminating native contrasts (/d/-/ð/) (Sundara, Polka, & Genesee, 2006), and 10- to 12-month-old English infants performed significantly better in discriminating native contrasts (/l/-/r/) compared to 6- to 8-month-old English infants (Kuhl et al., 2006).

Furthermore, the infants' capacity to discriminate non-native speech sound contrasts is supported by neural measures of brain activity during speech sound processing. Rivera-Gaxiola, Silva-Pereyra, and Kuhl (2005) reported that 7- and 11-month-old English infants show discriminatory evoked potential responses (ERPs) to native English consonant contrasts. Unlike ERPs of 7-month-old infants, ERPs of 11-month-old infants, as a group, did not reveal an evidence for discriminating non-native English consonant contrast. However, dividing the 11-month-old infants into two groups based on their individual ERP patterns (larger P150–250 and N250–550 discriminatory



responses), both groups of infants show significant discrimination of non-native contrast. Rivera-Gaxiola et al.' findings indicate an increase in neural responses to native contrasts and a diminished, but not lost, neural response to non-native contrast. Rivera-Gaxiola et al.'s findings are consistent with Cheour et al. (1998)'s findings of a decreased neural response for discriminating non-native contrasts. One year-old Finnish infants show small MMN amplitude for non-native (Estonian) vowel contrast compared to larger MMN amplitude observed on six-month-old Finnish infants and one-year old Estonian infants. Overall, infants' speech perception becomes tuned to ambient language input, and language-specific exposure facilitates their discrimination of native contrasts; however, infants do not fully lose the ability to discriminate non-native contrasts at an older age.

The developmental changes during infancy that are due to ambient language experience are well explained by the Native Language Magnet (NLM) theory which was proposed by Kuhl in early 1990s (Kuhl, 1994). This model posits that human adults perceive speech sounds categorically. According to the NLM, speech sounds are grouped into different phonetic categories with the best representations (prototypes) of the acoustic stimuli located in the center of the category. These prototypes—which emerge from the experience of one's native language—form perceptual magnets that warp the perceptual space around them, pulling all the adjacent stimuli toward the center and causing them to be perceived as belonging to the category. In other words, a listener may hear all the variants of a phoneme as being exemplars of that phoneme (i.e., prototype). In addition, the NLM accounts for the process of developing language-specific representations (prototypes) during infancy. Infants are born with a general auditory

mechanism that allows them to perceive sounds as universal phonetic categories, which are not human-specific, since animals also have such a perception pattern. With exposure to language-specific input, an infant retains the information about the distributional properties of sounds used in the ambient language. With additional language experience, they retain (store) information (representations) of the speech sounds of the ambient language, which reshapes the universal categories to become language-specific, and so some categorical boundaries are maintained and others disappear or reform. Thus, a prototype of each phonemic category evolves in the center of each category (Kuhl, 1994, 2004).

Given that infants' perceptual abilities are shaped by speech input during the first year of life, what is the mechanism by which an infant learns the native language phonetic categories? One part of the answer to such a question is related to infants' reliance on their statistical learning mechanism to decode the regularities in distributions of speech inputs. Maye, Werker, and Gerken (2002) studied 6- and 8-month-old infants' sensitivity to frequency and the probability distribution of speech sounds in the inputs. Infants were assigned to two groups: the first group was familiarized with more tokens from the two end points of the [da]-[ta] continuum (i.e., /ta/ and /da/) and a fewer tokens from the center of the [da]-[ta] continuum (bimodal group); the second group was familiarized with more tokens from the center of the [da]-[ta] continuum and a fewer tokens from the two end points of the [da]-[ta] continuum (unimodal group). In the testing phase, each infant listened to eight testing trials: four trials containing stimuli from the center of the [da]-[ta] continuum and four trials containing alternations of two

stimuli from the end of the [da]-[ta] continuum. Infants who were familiarized with stimuli from the two end points of the [da]-[ta] continuum (bimodal group) showed novelty preference for the testing trials containing stimuli from the center of the [da]-[ta] continuum, which suggests that they were able to signal the differences between the novel stimuli from the center of the [da]-[ta] continuum and the familiar stimuli from the two end points of the [da]-[ta] continuum. Infants who were familiarized with stimuli from the center of the [da]-[ta] continuum (unimodal group) showed no preference for the two types of testing stimuli. Maye et al.'s (2002) findings suggest that infants are sensitive to the frequency of the phonetic distribution in the input and are able to extract such information and use it to construct categorical phonetic representations of ambient speech sounds. Thus, statistical learning is an important mechanism that explains how infants derive and extract the distributional properties of speech inputs.

In summary, perceptual studies show that infants are born with an ability to discriminate different speech sounds. Language input exposure imposes a reorganization of infant perception with the shift to language-specific perception of vowels and consonants occurring between 6 and 12 months of age. Such effects from the ambient language cause infants and adults to perform well in discriminating native phonetic contrasts compared to non-native contrasts. However, the development of infant speech perception is a complex process. Infants come to the speech task with an auditory system that is still maturing with respect to its structures and functions. Although not covered in detail in the preceding review, infants also come to the speech task with initial perceptual biases; for example, a preference for corner vowels over non-corner vowels (Polka &

Bohn, 2003, 2011), for speech over non-speech (Vouloumanos & Werker, 2004), maternal voice over an unfamiliar female voice (Kisilevsky et al., 2003), and so on (summary in Gervain & Werker, 2008; Polka, Rvachew, & Mattock, 2007; Werker & Yeung, 2005). In addition, infants apply different cue weighting strategies for word segmentation, such as stress pattern (Jusczyk, Houston, & Newsome, 1999), allophonic cues, and syllable co-occurrence (Gervain & Werker, 2008; Werker & Yeung, 2005); suggestive of statistical learning (Maye, Weiss, & Aslin, 2008; Maye et al., 2002). As a result, infants' intake of speech input is a product of a complex interaction among different aspects of their auditory system and processes.

### **1.2.2 Development of speech production during infancy**

This section reviews the literature of the development of speech production during infancy. It begins with a discussion about prelinguistic vocalizations and then examines the development of speech sound production during infancy.

The experience of newborns with vocal production begins with crying at the moment of birth. From then until the production of their first meaningful word, their vocal production undergoes developmental changes throughout their first year of life. This prelinguistic period is commonly described in terms of overlapping stages of vocal development (Oller, 1980; Stark, 1980). The overlapping between stages is the norm, accounting for the infants' individual variations in progress rate through the stages (Oller & Lynch, 1992). Oller (1980) and Stark (1980) also provided compatible descriptions of the early vocal development of English infants that agree with subsequent studies on language other than English (e.g., Koopmans-van Beinum & van der Stelt, 1986, for

Dutch infants). Following are Oller's (1980) and Stark's (1980) combined descriptions of an infant's first year vocalizations:

(a) Stage One (the reflexive vocalization/phonation stage [birth – 1.5 months]): a newborn's vocalizations include crying, fussing, and vegetative sounds. These vegetative sounds are involuntary sounds related to respiration and nutrition—such as coughing, sucking, and swallowing—or to physical activity such as grunts and sighs (Vihman, 1996). Oller (1980) describes reflexive sounds (i.e., vegetative sounds) as non-arbitrary vocalizations elicited by internal or external stimulations, whereas non-reflexive sounds are formed by arbitrary sound pairing. Many of the sounds produced at this stage belong to a category that he calls *quasi-resonant nuclei*, which includes vowels or syllables with normal phonation and sound energy below 1200 Hz. These vowels and syllables are perceived as nasalized, and these sounds are produced with the mouth closed or near-closed due to the tendency of the tongue to rest high against the soft palate.

(b) Stage Two (the cooing and laughter/ primitive articulation stage [1.5 months to 4 months]): infants begin to produce pleasant and comfort sounds with less frequent crying. These comfort vocalizations are produced in relation to adult interactions (speaking, smiling, or face nodding to the infant) as an infant looks at interesting objects, or vocalizes after a pleasant feeding or to reflect excitement. At this stage, infants produce more vowels; and some consonant-vowel (CV) syllables with *fully resonant vowels* (sound energy above 1200 Hz), and

consonant-like sounds that include stops, velar-like, clicks, friction noises, and trills. In addition, the proportion of non-reflexive utterances increases.

(c) Stage Three (the vocal play/expansion stage [4 months to 7.5 months]): infants begin to produce different types of vocalizations related to playing and exploring the surrounding environment (e.g., toys, social interaction). Infants at this age also produce more fully resonant vowels with or without consonants. At this stage, vocalizations include raspberries (i.e., bilabial or labiodental trills or frication noises), squealing (i.e., normal phonation breaks at some point leading to falsetto), growling (i.e., low pitch, creaky voice), yelling, and ingressive-egressive sequences. The production of *marginal syllables* marks the end of this stage. These syllables are CV sequences, but their timing and articulatory aspects are not speech-like; however, at this stage, their proportion is still low.

(d) Stage Four (the reduplicated babbling/canonical babbling stage [6 months to 10 months]): it includes the production of a series of CV syllables, and the same syllables tend to be repeated (e.g., babababa). In this stage, infants produce speech-like CV syllables in terms of their articulatory and durational aspects known as *canonical syllables*. Parents perceive this type of babbling as their child talking or trying to talk; however, no meaning is associated with this type of babbling (Oller, 1980). Some infants also produce some non-reduplicated canonical syllables as frequently as they produce reduplicated syllables.

(e) Stage Five (the non-reduplicated babbling/variegated babbling stage [10 months to 14 months]): infants produce a greater variety of syllable shapes with

different consonants and vowel combinations (e.g., V, VC, CV, CVC). In addition, infants tend to produce a series of multisyllabic utterances with varied stress and intontional patterns that are perceived as “gibberish”. During the transition from babbling to meaningful speech, infants produce *protowords*, which according to Vihman’s (1996) definition are utterances that approximate word shapes but lack a form-meaning association (such as /dido/, and /dejo/). However, some scholars such as Menn (1983) view these utterances as meaningful, since they are bound to situations or contexts (such as *woof-woof* for *dog*), although they do not have adult phonetic, phonology, and/or semantic targets. The production of *protowords* marks the last step before the production of meaningful words around the beginning of the second year of life.

As infants progress through the prelinguistic and first word stages, they produce a great variety of speech sounds. With respect to vowel development, infants start with central, neutral vowels, and as they grow older, the vowel space expands toward its corners with the production of more peripheral vowels. Buhr (1980) and Kent and his associates (Kent & Bauer, 1985; Kent & Murray, 1982) conducted thorough investigations on vowel development during infancy. Buhr (1980) investigated the vowel space of one English-learning infant longitudinally from age 16-64 weeks. He based these investigations on a bi-weekly recording and identified vowels perceptually and acoustically. Initially, the infant’s most commonly produced vowel was the middle vowel ([ɛ]). Toward the age of 64 weeks, the infant used more peripheral vowels (mostly [ʊ], [æ], [ɪ], and [i]) along with middle vowels ([ɛ] and [ʌ]). Using longitudinal

investigations, other studies (Kent & Bauer, 1985; Kent & Murray, 1982) described vowel development in infants at different intervals of the first year of life. Kent and Bauer (1985) described the first formant frequency (F1) and the second formant frequency (F2) at 3, 6, 9, and 12 months. Even though the F1-F2 range increases with age, the predominant vowels are central and mid-front vowels. With respect to infants at 12 months of age, the central, low-mid front, and low-front vowels continue to dominate. In addition, the most frequent vowels in V and VV syllables are /ʌ/, /ə/, /ɛ/, /æ/, /e/, and /o/, which suggests the production of peripheral vowels along with central vowels at 12 months. Furthermore, Davis and MacNeilage (1995) described the vowels from the onset of canonical babbling (about 6-7 months of age) to three and half years of age. They confirmed that the most frequently produced vowels are central, mid, and low-front for the whole age range.

With respect to consonantal development, many scholars (e.g., Kent & Bauer, 1985; Locke, 1983) conducted intra-linguistic investigations of consonant acquisition, reporting that nasals, stops, labials, and glides tend to dominate in early vocalizations compared to other sound classes. For example, Kent and Bauer (1985) showed that the proportions of stops, glides nasals, gottals, and coronals are higher compared to other manner and place sound classes in the vocalizations of one-year-old English-learning infants (e.g., 74% stops, 5% glides, 10% nasals, and 11% fricatives). Similarly, some scholars (Ingram, 2008; Locke, 1983; Vihman, 2006) reviewed the data on segment acquisition during infancy within and across languages. Despite inter- and intra-linguistic variations, stops, labials, nasal, and glides are favoured, and tend to be acquired very early compared to other speech sounds. For example, Vihman's (2006) study analyzed



the early words produced by 33 18-month-old children learning one of six languages (American English, British English, Finnish, French, Italian, and Welsh), which showed that infants' phonetic repertoire consists mostly of stops, nasals, labials, and alveolars. Thus, infants have a limited phonetic repertoire in which the most frequent consonantal sounds are stops, labials, nasals, and glides while fricatives and liquids are often not found.

While infants produce a variety of speech sounds (vowels and consonants) and speech-like utterances, they also produce non-speech vocalizations. These non-speech utterances and many speech sounds are produced in a manner that is not adult-like, which imposes a challenge to the investigations of infant early vocalizations that use phonetic transcription. According to Oller (2000), human vocalizations can be categorized into four groups: (1) vegetative sounds that are related to respiration, swallowing, and digestion (e.g., coughing, sneezing, hiccupping), (2) fixed vocal signals that have communicative intents and are associated with desire (e.g., crying, laughing), (3) protophones that are pre-speech and have a similar sound quality to speech but are not adult-like except for canonical babbling; they may have social, practice, or play purposes but do not have linguistic meaning (e.g., quasivowels, gooing, fully resonant nuclei, marginal babbling, canonical babbling), and (4) speech (e.g., words, phrases, sentences). Oller criticizes traditional phonetic transcription and acoustic approaches to describing infant vocalizations or protophones as “shoe-horning them into mature alphabetical segment types” (2000, p. 60). According to Oller (2000), the use of the International Phonetic Alphabet (IPA) to study the sounds produced by young infants results in misleading conclusions about infant vocal development, such as the Jakobsonian

(Jakobson, 1941/1968) discontinuity view (i.e., babbling is random and not related to speech). Oller's (2000) criticism of the phonetic transcription approaches stems from the fact that infant vocalizations include speech behaviours that cannot be described accurately by the IPA as forms of "well-formed syllabic or segmental units" (p. 4). In addition, acoustic descriptions (e.g., duration, fundamental frequency, formant frequencies) of infant vocalizations do not consider the articulatory aspects of speech sounds and do not provide an understanding of the relationship between early vocalizations and speech, since many of an infant's sounds are not speech such as squeals, crying (Oller, 1986, 2000; Oller & Lynch, 1992). Further, according to Oller (2000), infants' vocalizations are variable in terms of specific phonetic categories (specific vowels or consonants), although these vocal patterns sustain a broad vocal category of protophones.

Rather than describe infant vocalizations by using acoustic analysis and phonetic transcriptions, Oller (1986, 2000) proposes an infraphonological approach. This model consists of the following elements: (1) two separate operational-level categories or units—one for adult speech (e.g., syllables, segments, features) and one for infant vocalizations (protophones), (2) infraphonological principles that "[specify] principles generating the entire class of potential well-formed operations units and ... [the] properties of utilization and function of such units" (Oller, 2000, p. 12), and (3) physical parameters that include acoustic features (e.g., durations, frequencies) and articulatory features (e.g., the movements and position of articulators). Although Oller (2000) does not clearly describe the exact roles of the infraphonological principles, they may function as hidden units that manipulate the physical parameters and connect them to operational

units. Infants' operational units (i.e., protophones) have different physical parameters from adults' operational units (e.g., segments). Therefore, Oller (1986, 2000) stresses that infant vocalizations should be described according to the infraphonological perspective, which is based on the acoustic features and articulatory parameters of protophones (e.g., quasivowels, gooing, marginal babbling, full vowels, raspberries, squealing, and canonical babbling) rather than adult units (e.g., segments). A further description of this infraphonological and phonation coding is presented in the Methods chapter.

### **1.2.3 Variables influencing infant speech production**

As infants progress through the pre-speech stages, many variables can influence the changes in their vocal outputs. Two of these variables are biological influences (maturation of the vocal tract and oral motor control) and linguistic environmental influences. However, the central controversy is how these two variables are involved in the process of babbling. In other words, the controversy is concerned with the age at which the changes in babbling are not merely based on biological influences but rather on a more complex interaction between the physiological and physical maturation of the vocal tract and the ambient language input. de Boysson-Bardies et al. (1989) illuminated this debate about the two hypotheses by pointing out that the *independence hypothesis* claims that the developmental changes in infant vocalizations are due to the articulatory mechanism and are independent from the perceptual system, whereas the *interactional hypothesis* claims that both motor and perceptual mechanisms are involved in the production of infant speech. The *interactional hypothesis* predicts the early effects of ambient language input on babbling and suggests that infants learning different languages

babble differently, whereas the *independence hypothesis* predicts a lack of early influences of language-specific exposure on babbling, and thus suggests that infants learning different languages babble similarly. A further discussion of these two variables follows.

### **1.2.3.1 Biological influences**

The biological effects on infant early speech outputs are due, in part, to the differences in the size and shape of infant vocal tract (VT) structures compared to adults. The course of the development of these VT structures passes through periods of growth and *anatomic restructuring*—a term used to describe the anatomic changes in the size and shape of the VT (Vorperian et al., 2005). Peterson-Falzone and Monoson (2003) illustrate the complexity of the structure in their description of the adult vocal tract: it is formed of 15 muscles, 6 bones (mandibular, maxillary, sphenoidal, occipital, temporal, and hyoid), and 4 cartilages (thyroid, arytenoid, corniculate, and cricoid). The length of the VT is 15 cm in adult females and 18 cm in adult males; and the length of the vocal fold is 18 mm in adult females and 12 mm in adult males. However, newborns have a small VT of 8 cm (on average) in length, small and short lips, a relatively large tongue almost filling their oral cavity, a shorter larynx of 6 cm (on average), a sloping angle of the oropharyngeal cavity (compared to the 90 degree angle between the adult oral cavity and the larynx), a velum and epiglottis close together, a high larynx position in the neck, short vocal folds of 4 mm (on average), accumulations of fat pads in the cheeks, and overall different sizes of VT structures compared to adults (Kent, 1992; Peterson-Falzone & Monoson, 2003; Thelen, 1991).

As infants mature, many anatomical changes occur in the VT structures, which include but are not limited to a lengthening of all the VT structures, a descending of the posterior part of the tongue, alterations in the intrinsic and extrinsic nature of the orofacial muscles, a forward and downward movement of the soft palate, less fat accumulation in the cheeks, and a descending pharynx (Kent, 1992; Peterson-Falzone & Monoson, 2003; Thelen, 1991). Several studies have examined the developmental changes in the VT anatomy by using techniques such as magnetic resonance imaging, videofluoroscopy, and X-ray reporting of the physical changes of the different portions of the vocal tract (Lieberman, McCarthy, Hiiemae, & Palmer, 2001; Rommel et al., 2003; Siebert, 1985; Vorperian, Kent, Gentry, & Yandell, 1999; Vorperian et al., 2005; Vorperian et al., 2009). Across these studies, infants show remarkable changes in the length, shape, and size of the VT and its structures; for example, the distance between the velopharyngeal valve to the tongue base increases from 20 mm in 3-month-old infants to 27 mm in 4-year-old children; the distance between the tongue base to the entrance of the larynx increases from 5 mm in 3-month-old infants to 9 mm in 4-year-old children (Rommel et al., 2003). In addition, infants exhibit a descending of the apex of the mandibular angle, hyoid bone, larynx, and vocal folds relative to the palate (Lieberman et al., 2001), whereas the maxilla and mandible thicken as age increases (Vorperian et al., 2005). The tongue size doubles many times from birth to adolescence, reaching maturity in late childhood, although in some individuals, the tongue may continue to grow a bit more (Fletcher & Daly, 1974; Siebert, 1985; Sperber, 2001). Autopsy examinations of lingual growth from 25 weeks of fetal age to 10.5 years on subjects with normal craniofacial structures indicate an increase in tongue size, which is about a 13 times

increase in weight from 2.6 grams to 33.4 grams, a 2.3 times increase in length from 2.6 cm to 6 cm, a 2.3 times increase in width from 1.9 cm to 4.5 cm, and a 2 times increase in thickness from 0.9 cm to 1.8 cm (Siebert, 1985). Furthermore, these changes in the VT pass through a period of accelerated growth from birth to 18 months during which many VT structures achieve between 55% to 80% of the adult size (Vorperian et al., 2005). These rapid changes in VT have significant implications on infant capacity to articulate speech sounds (a further discussion of these implications follows). Overall, an infant's VT undergoes structural changes in size, shape, and length that are rapid during the first 18 to 24 months of life.

An additional biological influence on infants' early vocalizations is the immaturity of oromotor control. Newborns have immature control of their lips, tongue, jaw, and velopharyngeal valving; and an undifferentiated jaw-lip coupling (Kent, 1992; Thelen, 1991). A few physiological studies use electromyographic recordings of some orofacial muscles involved in speech execution to provide developmental data about oromotor control (e.g., Green, Moore, Higashikawa, & Steeve, 2000; Green, Moore, & Reilly, 2002; C. A. Moore, Caulfield, & Green, 2001; C. A. Moore & Ruark, 1996; Steeve, Moore, Green, Reilly, & McMurtrey, 2008). These studies focus on investigating whether infants' oromotor control for speech is distinct from non-speech (e.g., chewing) and/or whether infant's oromotor control for speech is different compared to adults. First, the findings across these studies show that the oral motor control for speech is distinct from the oral motor control of the early primitive behaviours. For example, the coordination of the mandibular muscles (e.g., masseter, temporalis, agonist and antagonist muscle groups), as measured by the electromyographic (EMG) activities of 9-

and 15-month-old infants, was different during speech (babbling and spontaneous speech) compared to non-speech behaviours (chewing and sucking). Speech requires a high coupling (correlations) among the mandibular muscles compared to the lower coupling for non-speech behaviours (C. A. Moore & Ruark, 1996; Steeve et al., 2008). These differences replicated the findings from the adult data, which showed a task-specific coordination of the mandibular muscles for speech and non-speech (C. A. Moore, Smith, & Ringel, 1988). In addition, the differences in the level of muscle coupling are reported for other speech musculature; for example, the upper and lower lip muscles of two-year-old children show a different level of coupling (inter-correlation) during the production of bilabial speech sounds compared to the production non-speech movements such as chewing and lip protrusion (Ruark & Moore, 1997).

Second, the findings across the just-cited studies also show that infants' speech oromotor control is not mature at birth. For example, Steeve et al. (2008) compared the coordination of mandibular muscles at 9 months, 15 months, and in adults. Their findings show an increase in the muscle coupling and stability of the mandible across non-speech, babbling, and speech, from infancy to adult, which suggests a refinement of coordination as age increases. Such a finding of increasing oromotor control with age is consistent with other studies (Green et al., 2000; Green et al., 2002). These studies examine the development of jaw-lip control in the production of bilabial CVCV words. The kinematic measures were obtained for the upper lip, lower lip, and jaw from one-year-old, two-year-old, six-year-old, and adult subjects. The evidence showed that one-year-old infants show great jaw displacement during the oral closure for bilabial production, with a lack of contribution by the lower lip. Two-year-old infants showed an

increased contribution of the lower lip in the oral closure, with a slight decrease of jaw displacement (although the jaw still dominates the oral closure). Six year-old children began to show an adult-like pattern of oral closure in which both the lower lip and jaw contributed to the closure in a synchronized rising-falling movement, although this pattern showed more variation compared to adult subjects, with a minimum role played by the upper lip. Such a finding indicates that control over these articulators is gained gradually, and the control of certain articulators occurs before others (Green et al., 2000; Green et al., 2002). However, kinematic studies in development of speech motor control of other articulators during infancy (e.g., velopharyngeal valving, soft palate, and tongue) do not exist due to lack of appropriate technical tools (C. A. Moore, 2004). As a result, many scholars tend to infer their conclusions on the speech oromotor development during infancy based on indirect analyses such as vocal tract simulation (e.g., Menard, Schwartz, & Boe, 2004).

Consequently, the evidence suggests that anatomical and physiological variables constrain infants' capabilities to produce speech sounds, which limit their phonetic inventory. Based on earlier research by Sander (1972), Kent (1992) conducted a motoric analysis of the consonant acquisition of children learning English that showed that infants tend to first acquire consonants requiring less motor adjustment to produce (e.g., bilabials, nasals) compared to consonants requiring fine force coordination among articulators (e.g., fricatives). Such a finding is consistent with crosslinguistic findings that labials, stops, nasals, and glides occur frequently in the early consonantal inventory of infants acquiring different languages (Ingram, 2008; Locke, 1983; Vihman, 2006). In addition, Kent (1992) conducted a similar motoric analysis for the order of feature



acquisition based on the hierarchy of the five levels proposed by Dinnsen (Dinnsen, 1992; Dinnsen, Chin, Elbert, & Powell, 1990). The hierarchy levels are ordered from the least complex (fewer feature contrasts) to the most complex (more feature contrasts). The phonetic inventory for children learning different languages (Spanish, Quiche, and English) can be assigned, hence universal hierarchy (Dinnsen, 1992); however, some crosslinguistic differences have been observed in the order of the acquisition of some features. For example, the delayed-release feature occurs late in English compared to Japanese and Italian (Locke, 1983). According to Kent's (1992) motoric analysis, level A requires the regulation of the velopharyngeal valving and rate of articulatory movement for stop, nasal, and glide production; level B requires coordination of laryngeal and supralaryngeal timing for the production of the voicing distinction; level C requires a complex tongue configuration and fine force control for fricative and affricate production; level D requires tongue shaping with fine motor control for lateral and retroflex production; and level E requires more advanced tongue control for the production of "strident and non-strident" or "lateral and retroflex" distinction.

Furthermore, biological constraints limit infants' capability to produce vowels. An infant's vowel space is small and centered on neutral vowels (/ə/) due to their small vocal tract and immature oromotor control. Gradually, as the vocal tract increases in size and articulatory control matures, an infant's vowel space expands toward its corners (/i/, /u/, /æ/, /a/), for example, as described by Buhr (1980) and Kent and his associates (Kent & Bauer, 1985; Kent & Murray, 1982). In addition, Menard et al. (2004) created a simulation of the vocal tract and acoustic vowel space of French-speaking subjects at

birth, 4 months, 10 months, 16 months, and 21 years of age by using an acoustic-to-articulatory model (a Variable Linear Articulatory Model). Their results showed that the infant vocal tract is capable of producing all 10 French vowels; although due to its small configurations, it tends to favour the production of front and low vowels. Similarly, acoustic investigations of 10- to 18-month-old English and French infants revealed a similar pattern of findings in which young infants showed restricted vowel space, and as age increased, the center of the vowel space expanded toward the peripheral corners, particularly toward the high-front and high-back corners (Rvachew et al., 2006). Thus, infants produce the speech sounds that are phonetically permissible by their articulatory systems.

In addition, biological constraints extend to infants' early word forms. During the second half of the first year, infants start producing stereotypical rhythmic jaw movements combined with phonation producing syllabic sequences with adult-like timing, which are canonical syllables. After infants gain the capability to produce these syllabic sequences, they appear to focus on the production of whole-word units/forms (templates) to match adult target words, but in patterns that are phonetically accessible to them (Vihman & Croft, 2007). The forms are simple syllable shapes, and many are consistent with the Frame/Content Theory's assumption (MacNeilage & Davis, 2000, 2001) that the earliest syllable shape is a CV that is the result of phonation with alternative opening-closing cycles of the mouth, with a minimum contribution by the articulators. The mouth opening is for vowel production, and the mouth closing—to different degrees—is for consonant production. *Frame* refers to the syllable structure, whereas *content* refers to segments. The early CV-combinations are usually the

following: (1) labial consonants (e.g., /p/, /b/, /m/) with central vowel, (2) coronal or tongue front consonants (e.g., /t/, /d/, /n/) with front vowel, and (3) dorsal (e.g., /k/, /g/) with back vowel. Moreover, for early stages of speech production, Frame/Content Theory emphasizes the major role of jaw involvement with a lesser involvement of the tongue and lips. Such suggestion agrees with kinematic investigations of the development of jaw-lip control that found that infants gain control over the jaw first and then lips and that the jaw dominates the closure for bilabial production (Green et al., 2000; Green et al., 2002).

In addition, Frame/Content Theory assumes that “the mandibular cycle underlying the speech frame may have been exapted (‘borrowed’) from the mandibular cycle originally evolved for mammalian ingestive functions—chewing, sucking, licking” (MacNeilage & Davis, 2001, p. 696), suggesting that speech emerges from early primitive behaviors, and hence the early lack of task specificity for the muscles involved in both speech and non-speech functions. Such an idea is under reconsideration, since some scholars have found evidence for speech specific coordination (C. A. Moore & Ruark, 1996; Steeve et al., 2008; Wilson, Green, Yunusova, & Moore, 2008); as young as nine month old (Steeve et al., 2008). Speech requires rapid and independent movements of the articulators (e.g., lips and tongue), whereas sucking, for example, requires vertical tongue movement coupled with jaw movement (Wilson et al., 2008).

Furthermore, the anatomical and physiological changes in the vocal tract demand a constant change in the control of the articulators (Vorperian et al., 2005; Vorperian et al., 2009). Infants are confronted with the challenge to continue to produce the target

sounds as their articulatory systems change, particularly during growth spurts from birth to 18 months. To some extent, they are able to modify their speech production while undergoing these developmental changes. For example, Callan, Kent, Guenther, and Vorperian (2000) explain how infants continue to produce vowels during the developmental changes in their vocal tract by applying a Maeda articulatory synthesis, which is part of the Directions into Velocities of Articulators (DIVA) neural network computational model (Guenther, 2006). By emphasizing the importance of auditory and somatosensory feedback to speech production, the DIVA model provides an explanation of how the infant, during growth, maintains sufficient control over the articulators to produce the target vowel categories despite dramatic reshaping of the vocal tract. The DIVA model further suggests that the semi-random movements of the articulators during early infant vocalization generates auditory, tactile, and proprioceptive feedback that enables a development of an internal model of the mapping between acoustic goals and the vocal tract shapes that can produce the acoustic targets. The acquisition of stable acoustic goals early in life allows the infant to adapt to developmental changes in vocal tract morphology by updating the internal model as needed, which allows the system to produce speech with stable acoustic-phonetic characteristics while the vocal tract is undergoing a series of changes.

### **1.2.3.2 Linguistic environmental influences**

Many aspects of the infant's environment—such as visual information (Kuhl & Meltzoff, 1996) and social interactions (Goldstein & Schwade, 2008)—may contribute to changes in speech production. However, most research has focused on the role of speech

input as received by the infant through the auditory modality. The nature of the speech inputs and the infant's ability to access and process them are significant for normal speech development (Oller & Eilers, 1988; Oller, Eilers, Bull, & Carney, 1985).

One line of research on the role of auditory input on normal speech production focuses on the effects of hearing impairment on infant babbling. Several studies have shown a good consensus on the differences in the vocalizations between hearing-impaired infants and normal-hearing infants due to the lack of or poor auditory input (Eilers & Oller, 1994; Oller & Eilers, 1988; Oller et al., 1985; Stoel-Gammon, 1988; Stoel-Gammon & Otomo, 1986). Hearing-impaired infants with severe and profound hearing loss started to produce canonical babbling after 10 months of age (between 11 to 25 months), whereas hearing infants produced canonical babbling between the ages of 6-10 months. Moreover, none of the hearing-impaired infants met the minimum criterion for the canonical babbling stage which is a ratio of 0.2 canonical syllables to utterances (Oller & Eilers, 1988; Oller et al., 1985). In addition, hearing-impaired infants showed a small consonantal repertoire with little or no change with age and a great proportion of labials over alveolars due to the presence of visual cues. They also produced fewer multisyllabic utterances compared to hearing infants (Stoel-Gammon, 1988; Stoel-Gammon & Otomo, 1986). In addition, Eilers and Oller (1994) reported that canonical babbling correlates with the age at which hearing aids were fitted. This finding indicates that speech production (in this case, canonical babbling) occurs when the hearing-impaired infants have access to speech input. Further, a mild conductive hearing loss as a result *otitis media* (OM) influences infants' vocalizations. For example, Rvachew and her associates (Rvachew, Slawinski, Williams, & Green, 1996; Rvachew, Slawinski,

Williams, & Green, 1999) showed that 6- to 18-month-old infants with early-onset-OM (prior to age 6 months) have restricted vowel space, fewer utterances with canonical syllables, more utterances with precanonical syllables, and a low ratio of canonical syllables compared to their counterpart infants with late-onset-OM (between 6-18 months). Therefore, hearing impairment restricts an infant's ability to access speech input, which imposes negative effects on speech production.

Another line of research on the role of auditory input on normal speech production focuses on the effects of the linguistic environment on infant early speech production. The search for such influences is motivated by several notions. Most world languages show differences in their phonetics and phonology, and infants are exposed to their ambient language from birth (or even during pregnancy) through what is known as infant-directed speech (IDS). The IDS is characterized by a high fundamental frequency, a slow rate, exaggerated intonation contour, stretched vowel spaces, and frequent pauses (Fernald, 1985; Kuhl, Andruski, Chistovich, Chistovich, & et al., 1997). Despite the similarity in its characteristics, the nature of IDS varies across languages (Kuhl et al., 1997; Lee, Davis, & Macneilage, 2010). Kuhl et al. (1997) examined the vowels in IDS produced by 30 native American English-, Swedish-, and Russian-speaking mothers to their infants who ranged in age from 2 to 5 months. The mothers were asked to produce words containing point vowels (/i/, /a/, /u/) while (1) interacting with a native adult speaker (adult-directed speech [ADS]) and (2) interacting with their infants (IDS). The acoustic analysis based on the F1 and the F2 revealed that the vowel space was greater in the IDS compared to the ADS, hence, the exaggerated articulation of the vowels. In addition, crosslinguistic differences occurred in the production of vowels in the IDS

compared to ADS; for example, English mothers increased F2 in /i/, decreased F2 in /u/, and increased F1 and F2 in /a/; whereas Swedish mothers increased F2 and decreased F1 in /i/, decreased F1 in /u/, and increased F1 and F2 in /a/. Similar crosslinguistic differences in IDS were reported in the proportion of consonants and vowels. For example, Lee et al. (2010) examined the consonants and vowels in the IDS produced by 10 Korean and English mothers of 12-month-old infants. Their results showed significant crosslinguistic differences in sound classes with respect to the manner of articulation (stops, fricatives, affricates, and nasals), place of articulation (coronal, dorsal, and glottals), and vowel categories (e.g., low-central, high-front). Furthermore, infants prefer to listen to IDS over ADS. For example, 4-month-old infants prefer to listen to an 8-second-IDS sample over an 8-second-ADS sample (Fernald, 1985). Thus, infants' experience with ambient language starts very early from the last trimester of pregnancy (e.g., Kisilevsky et al., 2003), and IDS presents infants with language-specific input in a form that attracts their attention.

In addition, other lines of research exist that show a potential for infant babbling to be influenced by their language exposure. The research on the development of speech perception during infancy shows that infants are born with perception skills that are relatively more mature compared to their production skills. For example, infants are able to discriminate native and non-native speech sound contrasts. Infants' perceptual skills enable them to access and process phonetic information from the speech inputs in the ambient language. The effects of the phonetic properties of the ambient language in the form of IDS or even ADS are noted in infant speech perception during the first year of

life. As discussed previously, language-specific perception is developed as early as 6 months for vowels, and from 8 to 10 months for consonants (Kuhl et al., 1992; Werker & Tees, 2002). Moreover, infants change their vocalizations to match an adult target following short exposure, suggesting the effects of specific auditory input on vocalization. For example, infants who heard a 5-minute segment of /i/ produced significantly greater proportions of /i/-like vowels (Kuhl & Meltzoff, 1996 [for a review of this research, see the “Discussion” section]). Therefore, one may expect to observe differences in the babbling of infants exposed to different linguistic background environments before the first year of life.

However, previous research on the influence of the linguistic environment on babbling yielded two counter conclusions. Some studies failed to find significant differences in the babbling of monolingual infants exposed to different languages, whereas other studies showed crosslinguistic differences in babbling. A sample of these studies is reviewed in the following paragraphs—first the studies that showed no evidence and then the studies that showed positive evidence.

The findings from studies that provide no evidence for the presence of crosslinguistic differences in babbling are consistent with the *independence hypothesis* (de Boysson-Bardies et al., 1989) that suggests that the ambient language influences do not occur in early babbling, and that changes on babbling are determined by the maturation of the articulatory structures and functions until 18-24 months of age, as suggested by the frame-content theory (MacNeilage & Davis, 2000, 2001). By using previously published studies, Locke (1983) reviewed the consonantal inventory of 0- to



15-month-old infants learning one of 15 languages (age was not constant across the studies; for example, the English data included infants with ages ranging from 1-15 months, while the Thai data included infants with ages ranging from 10-11 months). Locke reported that a great similarity existed in the consonantal inventories across the 15 linguistic environments and concluded that: “The cross-linguistic studies of babbling seem to have no appreciable evidence in support of the babbling drift hypothesis” (p. 26). Locke’s conclusion may have resulted from his reliance on studies that used phonetic transcription as the method of identifying the consonants produced during the precanonical and canonical stages. Due to the immaturity of their articulatory system, infants do not master an accurate production of many consonants during the prelinguistic period. According to Locke’s review, only [m] and [b] were produced in babbling across the 15 languages. Locke’s methods (i.e., reviewing phonetic transcription studies and including precanonical vocalizations) may hinder the efforts to detect an effect of early language input because many precanonical vocalizations are produced during the first half year of life making IPA transcription difficult and less reliable, since adult coding may “shoehorn” infant vocalizations into mature forms as described by Oller (2000).

Subsequent researchers avoided transcribing precanonical sounds, and focused instead on transcribing sounds produced in speech-like utterances (e.g., canonical syllables), but still failed to find an effect of the ambient language environment on the phonetic characteristics of infant babble. As discussed previously, based on earlier research done in their laboratories, MacNeilage and Davis (MacNeilage & Davis, 2000, 2001) reported a similarity in the intrasyllabic organization of early utterances (including babble and first words) produced by infants learning different languages. They found that

the simple CV form (composed of motorically easier consonants) is the most common early syllable shape across world languages. Furthermore, they reported that certain consonants and vowels tended to co-occur within these syllables across languages (specifically, labials with central vowels, coronals with front vowels, and dorsals with back vowels), suggesting that the content of infant utterances is determined by a biomechanically motivated frame of cyclical jaw movements carrying an essentially static tongue. Similarly, Oller and Eilers (1982) used the phonetic transcription of canonical utterances to show that a great similarity existed between the babbling and early speech of 8 English- and 8 Spanish-infants (mean age of one year) in their consonantal and vowel inventories (e.g., both groups preferred to produce obstruents, especially unaspirated stops, and front or central vowels, in particular [i], [e], [ɛ], and [æ]). However, the findings of these two studies may not provide strong evidence against the presence of linguistic environmental influences on babbling. In contrast, the MacNeilage and Davis's (2000, 2001) studies aimed to show how the biomechanics of the mouth opening and closing imposed universal patterns on the early utterances of infants; but they did not claim that Frame/Content Theory provided evidence against the early ambient language influences on babbling. In addition, Oller and Eilers (1982) commented on their own findings as not suggesting "NO differences in babbling... but it does suggest that differences may be hard to find in the light of overwhelming similarities and rare production of non-universal phonetic elements" (p. 575). Further, Oller and Eilers recruited infants for their study from the Miami community where it is common for Spanish-speaking families to be bilingual. This fact is confirmed by remarks in the study that indicate that some of the infants' Spanish-speaking parents were competent in

English. In addition, their analysis is limited to recordings obtained at the age of one year, even though the authors indicated that all the infants in this study were part of another longitudinal investigation. These issues in the Oller and Eilers (1982) and MacNeilage and Davis (2000, 2001) studies may affect the generalizability of their findings.

Further evidence against the ambient language influences on babbling is reported by perceptual studies in which native speakers judged whether infant speech sounded like their native languages (Atkinson, MacWhinney, & Stoel, 1968; Engstrand, Williams, & Lacerda, 2003; Thevenin, Eilers, Oller, & Lavoie, 1985). The evidence reported across these studies was that native listeners failed to identify the infants' ambient language. For example, an earlier study by Atkinson et al. (1968) reported a failure of native English speakers to distinguish English speech samples from Chinese and Russian speech samples from infants who were 6-7, 10-11, and 16-17 months of age. However, Atkinson et al. did not claim that their findings rule out ambient language influences on babbling, since their study has many limitations, such as short speech samples (15 seconds), fewer speech samples (e.g., the identification test included only 2 speech samples from each language group at each age group), and a lack of statistical analyses. Opponents to the early role of ambient language input on infant vocal output tend to cite the failure of listeners in this study to detect language differences in babbling as evidence against the hypothesis of early ambient language influences without considering the methodological flaws of the study (e.g., Oller & Eilers, 1998).

Subsequent listening studies were conducted to examine language effects. For example, Engstrand et al. (2003) conducted listening studies in which English listeners

with phonetic experience listened to utterances produced by 12- and 18-month-old English and Swedish infants. Their task was to decide whether the utterances were produced by English or Swedish infants. The results showed that listeners tended to judge English utterances as produced by English infants, and Swedish utterances as produced by Swedish infants. However, the analysis of variance failed to reach significance, even in the case of the utterances produced by the 18-month-old infants. When the utterances included babbling, words, and imitations of mother's speech, a language effect was observed at both age groups (12- and 18-month-old English and Swedish infants). However, Engstrand et al.'s (2003) study was limited to only two age groups, and three out of five listeners were bilingual English and Swedish speakers and one listener was multilingual—English, Swedish and Estonian—characteristics that may have affected their responses when judging the babbling utterances. Moreover, the listening task occurred over six hours and was divided into one-hour sessions with no breaks, which is a long time for a listening study requiring listeners to be alert, and so fatigue may have affected the reliability of the listeners' judgments. In a similar study, Thevenin et al. (1985) also reported that 20 monolingual English and bilingual English and Spanish speakers failed to identify the language background of the utterances produced by English and Spanish infants of two age groups (7-10 months and 11-14 months). However, the authors reported a great individual variability in the listeners' judgments in that some were able to correctly identify the language background. In addition, the stimuli consisted of eight utterances from each infant, which is a small number to be a representative sample of infant speech. Overall, the findings of the studies providing

negative results against the early effects of ambient language input on babbling have some methodological issues that limit their conclusions.

In contrast to the studies that reported negative results, other studies provided positive evidence for the crosslinguistic differences in infant vocal output. Their findings support the influence of the linguistic environment on babbling and are consistent with the *interactional hypothesis* (de Boysson-Bardies et al., 1989), which suggests that changes in babbling are a product of an interaction among several variables including the biological variables (small vocal tract and immature speech motor control) and ambient language input. Some scholars have conducted listening experiments to verify the presence of ambient language effects. For example, de Boysson-Bardies, Sagart, and Durand (1984) investigated the effects of the linguistic environment on speech prosody. These researchers collected speech samples from Arabic-, Cantonese-, and French-speaking infants at 6, 8, and 10 months of age. With respect to perceptual judgements, from each infant, they took a 15-second speech sample that included marginal and canonical babbling. The French sample was paired with either an Arabic sample or a Cantonese sample. The findings showed that French speakers were able to identify the French sample most of the time for the French-Arabic pairs (75.8 % for 8 months and 74.4 % for 10 months), for the French-Cantonese pairs (69.4 % only for 8 months), and French-Arabic pairs (68 % for 6 months but by experienced phoneticians). In addition, listeners were able to identify the French sample from the Arabic sample of the 10-month-old infants when the samples had good intonation patterns, but failed to do so when the samples had poor intonation patterns. Thus, the findings suggested that early ambient language affected prosodic features (e.g., pitch and intensity contour) as young

as 8 months. However, de Boysson-Bardies et al.'s (1984) study included samples from only eight infants (two Arabic, two Chinese, and four French), and so its small sample size compromises the generalization of its findings across languages, since intra-linguistic differences of phonological development are very common (Ferguson & Farwell, 1975).

In different types of studies, ambient language effects have been reported based on data from the phonetic transcription studies of infant speech. These investigations revealed both universal patterns and language-specific differences (de Boysson-Bardies & Vihman, 1991; Lee et al., 2010; Levitt & Utman, 1992). For example, across languages, infants produced more labials, stops, nasals, and glides compared to fewer fricatives, affricates, and liquids. Infants also preferred to produce mid-front, mid-central, and low-central vowels. However, over and above these similarities, crosslinguistic differences were reported. For example, de Boysson-Bardies and Vihman (1991) analyzed the segmental inventories of 20 infants learning either English, French, Japanese, or Swedish from 9 months of age to the production of the first 25 words. Their analysis of the overall consonantal distributions in babbling showed the presence of crosslinguistic differences in the place of articulation. French and English infants produced significantly more labials (French infants showed greater labial production compared to English infants); French infants produced fewer velars (the difference was only significant between French and Japanese infants); and Japanese and Swedish infants produced more dentals (the difference was only significant between Swedish and French infants). In addition, crosslinguistic differences were reported with respect to the manner of articulation. French infants produced significantly fewer stops compared to the other

groups, and Swedish infants produced significantly fewer nasals compared to the other groups. According to the authors, these differences were in the same direction as predicted by the adult language based on their analysis of the consonantal distributions in the phonetic inventories of the adult target words. The distribution of labials was higher in French and then English; French had fewer velars; Japanese had higher dental distribution (English and Swedish were similar); French had a smaller percentage of stops; and Swedish had less nasal distributions. Furthermore, Levitt and Utman (1992) compared the babbling utterances of one French-learning infant and one English-learning infant at 5, 8, 11, and 14 months of age. Across the four age intervals, their findings showed that the English infant produced seven additional consonants (mostly fricatives and affricates) that the French infant did not produce. In addition, the English infant produced more stops, fricatives, affricates, and nasals compared to the French infant who produced more approximants. The English infant also produced a higher proportion of closed syllables than open syllables, whereas the French infant produced an equal amount of closed and open syllables. Levitt and Utman's (1992) described their findings of crosslinguistic differences in the infants as being similar to those observed in the adult language; specifically, English has phonemic affricates whereas French does not and more closed syllables are produced in English compared to French. Moreover, Lee et al. (2010) analyzed segmental distributions in the speech produced by 6 English and 6 Korean infants from age 8 to 12 months. They also noted some crosslinguistic differences in which the Korean infants produced more nasals, and the English infants produced more fricatives and glides. In addition, the English infants produced more glides, although both language groups produced few glides. Further, the Korean infants

produced a lot of low-central vowels and fewer low-front vowels, whereas the English infants produced more high-front, high-back, mid-front, and low-front vowels. Lee et al. (2010) commented on their findings as reflecting some differences in the segmental distributions in the IDS. For example, nasals were the second most frequent sound in the Korean IDS and higher in proportion compared to the English IDS; fricatives were higher in proportion in the English IDS compared to the Korean IDS; and high-front vowels were higher in the English IDS, whereas low-central vowels were higher in the Korean IDS.

Nonetheless, de Boysson-Bardies and Vihman's (1991), Levitt and Utman's (1992), and Lee et al.'s (2010) findings have some limitations. The Levitt and Utman study is based on case study investigations, and the crosslinguistic differences in the transcription data were not submitted to tests of statistical significance, except for the proportion of closed and open syllables. While de Boysson-Bardies and Vihman's (1991) study showed interesting findings about crosslinguistic differences with respect to the consonantal distributions in babbling and first words, the significant differences in babbling were not substantial (e.g., only in the distribution of labials among all four groups of infants). Further, Lee et al.'s (2010) study showed the presence of some crosslinguistic influences on the proportion of segmental inventories in the babble of Korean and English infants; however, the sample size is small, and a statistical crosslinguistic difference was found only between the frequency of occurrences of low-front and high-front vowels.



Further evidence for the early effects of ambient language on babbling is provided by crosslinguistic comparisons based on acoustic analyses of different aspects of infant vocalizations. For example, de Boysson-Bardies, Sagart, Halle, and Durand (1986) studied the long term spectrum (LTS) patterns obtained from adults and 10-month-old infants in three different language groups (French, Arabic, and Cantonese). The LTS gives a general indication of the spectral characteristics of the speech sample through showing the distribution of energy across speech frequencies (Kent & Read, 2002). The researchers calculated the LTS patterns from 40-second utterances. Their findings showed that language-specific LTS patterns exist. That is, infants learning the same language tend to have similar LTS patterns and so do adults. In addition, a similarity exists in the gross LTS patterns of adults and infants of the same language group. For example, French infants have single-peak LTS at 825 Hz, and French male adults have single-peak LTS at 400 Hz; whereas Arabic infants have a two-peak LTS at 830 and 1600 Hz, and Arabic male adults have a two-peak LTS at 225-750 Hz and 1600 Hz. Although de Boysson-Bardies et al.'s (1986) study showed a similarity between the infant and adult LTS patterns of the same language group, their findings are not based on any statistical comparison, and it is not clear how the LTS measurement provided information about the phonetic similarity between adult's and infant's utterances. In addition, the sample size of the infant participants is small (five to six infants in each language group).

In different acoustic investigations, Levitt and Wang (1991) and Hallé, de Boysson-Bardies, and Vihman (1991) reported the presence of some crosslinguistic differences on infant prosody. In their longitudinal study, Levitt and Wang (1991)

compared the rhythmic properties of the babbling from five 4- to 17-month-old English- and French-learning infants. Their acoustic analyses of the syllables produced in reduplicative utterances showed that final-syllable lengthening was longer in the French group compared to the English group. The French group also had more regular timing (less variability) of the non-final syllable compared to the English group. According to the authors, this crosslinguistic difference can be explained because French is a syllable-timed language in which the lengthening of the final-syllable is more salient, and non-final syllables have equal stress compared to the variable stress patterns of the English language, which makes syllables vary in length. This finding also was reported by Levitt and Utman (1992) in their analysis of syllable durations in the babbling of one English infant and one French infant. Further, the French group had longer utterances with more syllables per utterance compared to the English group; this finding may also be due to the crosslinguistic differences in the stress patterns in adult speech (Levitt & Wang, 1991). In a cross-sectional prosodic investigations, Hallé et al. (1991) showed that 18-month-old French infants tended to produced disyllables with rising contour, whereas the 18-month-old Japanese infants tended to produced disyllables with falling contour. In addition, the French infants tended to lengthen the final syllables compared to the Japanese infants. According to these authors, these crosslinguistic differences also are featured in the adult languages. Similar to the previously reviewed studies, the findings from Levitt and Wang (1991) and Hallé et al. (1991) investigating the presence of ambient language effects may not be conclusive. Levitt and Wang did not specify an approximate age for the emergence of such effects, since they reported on the language effects across a wide age range (i.e., 4 to 17 months), whereas Hallé et al. showed crosslinguistic variations at 18 months of age.

Ambient language effects may have occurred around 17 months of age, which is the age that is less controversial in the literature because this is the approximate time when toddlers begin to produce meaningful speech of one- to two-word utterances (de Boysson-Bardies, 2001).

Acoustic analysis also has been used to search for crosslinguistic differences related to the characteristics of infant vowels. In a frequently cited study that provides evidence for the ambient language effects on babbling, de Boysson-Bardies et al. (1989) collected speech samples from five 10-month-old infants in four different language groups (English, French, Arabic, and Cantonese) and found a significant language effect in the mean of the first formant frequency ( $M_{F1}$ ) and in the mean of the second formant frequency ( $M_{F2}$ ). In addition, these authors reported crosslinguistic differences associated with the location of the center of the infant vowel space derived from F1-F2. Unlike de Boysson-Bardies et al.'s (1989) study, Levitt and Utman (1992) examined a larger age range regarding the ambient language effects on center of the infant vowel space, although this study was somewhat limited in another respect because it focused only on one child for each language. Specifically, Levitt and Utman analyzed the  $M_{F1}$  and  $M_{F2}$  of one English infant and one French infant at 5, 8, 11, and 14 months of age. Their findings showed that  $M_{F1}$  was stable for both infants with a slight increase for the English infant and a slight decrease for the French infant. The English  $M_{F2}$  was stable until 11 months and then increased, whereas the French  $M_{F2}$  was stable until 14 months and then decreased. However, the just-cited studies (de Boysson-Bardies et al., 1989; Levitt & Utman, 1992) have some limitations that affected the level of support they provided for the ambient language effects on infant vowel space. For example, both studies examined

crosslinguistic differences only on the average vowel, i.e., the centre of the vowel space rather than an actual vowel produced by the infants. The average vowel covers a relatively large portion of the vowel space, since the study by de Boysson-Bardies et al. (1989) reported F1 values that ranged from 400 to 1600 Hz, and F2 values that ranged from 1250 to 3800 Hz. Further, both studies (de Boysson-Bardies et al., 1989; Levitt & Utman, 1992) did not describe how the groups of infants differed in their use of specific vowels.

Thus far, studies of the crosslinguistic differences in the infant vowel space have examined only the center vowel. This limited focus has occurred because an infant cannot be induced to produce a specific vowel, and thus the procedure that normally would be used to study differences in the adult production of vowels across languages cannot be replicated. For example, Escudero and Polka (2003) examined the acoustic characteristics of the specific vowels produced by adult speakers of English and French in Montreal. They asked these speakers to produce all the high and front vowels used in their language (English: /i/, /ɪ/, /u/, /ʊ/, /ɛ/ and /æ/; and French: /i/, /ɪ/, /y/, /Y/, /u/, and /ʊ/). They found numerous differences, especially regarding the high vowels. For example, all French high vowels had lower F1 values compared to English high vowels. In particular, the French /u/ in adult speech was more grave than the English /u/ due to lower values of F1. It is not possible to replicate this procedure with infants. However, Rvachew et al. (2006) introduced a means of studying the corner vowels in the infant vowel space of 10- to 18-month-old infants learning either Canadian English or Canadian French. Especially, the method involves reversing the procedure used in adult research by using the raw acoustic parameters to identify the vowels with the most extreme values of the features acute,

grave, diffuse and compact; and subsequently, submitting these vowels to various analyses. Rvachew et al. (2006) demonstrated that infants in both language groups expanded their vowel spaces toward the diffuse (high-front or /i/-like) corner and grave (high-back or /u/-like) corner, although French infants showed somewhat greater expansion toward the diffuse corner, and English infants showed somewhat greater expansion toward the grave corner. In addition, both groups showed smaller values of acuteness with age (i.e., maximum value of  $[F1+F2]/2$ ). Significant age-by-language group interactions were observed for the center of the vowel space as well: in the English-learning infants, a decreased  $M_{F2}$  (but a nonsignificant decline in  $M_{F1}$ ); and in the French-learning infants, a decreased  $M_{F1}$  (but an unchanged  $M_{F2}$ ). Furthermore, vocal tract modeling was used to show that the reported age-related changes in the corners and center of the vowel space were not a simple reflection of growth in the vocal tract. For some parameters, the changes were greater than would have been expected with vocal tract growth, and in others, the changes were less than would have been expected as a consequence of vocal tract growth. Therefore, both sets of changes (expansion into the diffuse and grave corners and restriction of acuteness values) appeared to reflect improvements in speech motor control; and  $M_{F2}$  changes appeared to reflect the differences in the English and French vowel inventories.

To summarize the literature on the infant vowel space reviewed thus far, several studies have described differences in the acoustic characteristics of infant vowels produced by infants learning different languages. In general, these findings have been interpreted within the framework of *the babbling drift hypothesis* (Brown, 1958). The hypothesis is that an infant will attempt to replicate the adult input—that is, it is expected

that the infant's vowel space will gradually approximate the characteristics of the adult vowel space in a linear fashion. In other words, the crosslinguistic differences in the acoustic characteristics of infant vocal output should parallel the crosslinguistic differences in adult vocal output. However, it is not clear whether the developmental process is as simple as this model would suggest. Furthermore, the *interactional hypothesis* (de Boysson-Bardies et al., 1989), which takes into account the infant's perceptual biases and skills in interaction with the limitations of the infant's motor control system and the input, does not predict a direct correspondence between infant vocal output and adult input. This principle is demonstrated in study by Rvachew, Alhaidary, Mattock, and Polka (2008). In the Rvachew et al. (2008) study, I obtained listeners' judgments from English and French adult speakers to determine: (1) the age at which expansion into the diffuse, grave, acute, and compact corners (/i/-, /u/-, /æ/-, and /a/-like corners, respectively) of the vowel space yields vowels that are perceptually accurate corner vowels; and (2) the ambient language effect on corner vowels. I analyzed speech samples collected from Canadian English and Canadian French infants aged 8-18 months. I used F1 and F2 values to calculate the feature values along the compact-diffuse ( $F2 - F1$ ) and grave-acute ( $[F1 + F2]/2$ ) dimensions for each infant. Vowels with minimum and maximum 10% values on the compact-diffuse dimension (/i/, /a/, respectively) and vowels with minimum and maximum 10% values on the grave-acute dimension (/u/, /æ/, respectively) were submitted to native English and French speakers for perceptual judgment. The findings showed the following. First, vowels with extreme values were perceived most of the time as corner vowels. That is, vowels with extreme diffuseness,

graveness, acuteness, and compactness values were judged to be /i/-, /u/-, /æ/-, and /a/-like vowels, respectively. Second, significant language effects occurred when listening to extreme grave vowels. Both listener groups heard more /u/ in the English sample compared to the French sample (i.e., English-learning infants produced more /u/ vowels compared to the French-learning infants). The language effects were noted before 12-months of age (12 % of the French grave vowels were judged to be /u/, while 30 % of the English grave vowels were judged to be /u/). This prominent language effect for /u/ (i.e., the occurrence of more /u/ in the English sample) was explained by the fact that the front area of the French vowel space is crowded (more complex) compared to English vowel space; and that the production of /u/ requires a precise articulatory configuration—a small articulatory variation in the production of /u/ (moving away the constriction from the velopalatal region in the infant vocal tract) may result in the production of a different vowel (Menard et al., 2004). For English infants, any variation in the production of /u/ may result in a vowel that may not sound like any English vowel, whereas for the French infants, any variation in the production of /u/ may result in a vowel that falls into another French vowel category. As a result, /u/ appears to be a strong attractor for English infants, which is reflected in their babbling because more /u/ was observed.

Based on the predications of the *babbling drift hypothesis*, we would expect to find French infants producing more /u/ compared to English infants (which we did not find), since English adult /u/ is less grave compared to French adult /u/, as is suggested by Escudero and Polka's (2003) study. If a researcher was looking for a drift in the infant vowel space toward adult language (i.e., infants trying to replicate adult input), the

findings in Rvachew et al.'s (2008) study of English infants producing more /u/ would not support this *babbling drift hypothesis*. According to this study, English and French infants did not replicate the adult vowel space; rather, their vowel spaces were influenced by the complexity of the vowel space of their ambient language. The English vowel space is less complex compared to the French vowel space, since French has more vowels, especially at the front area of the vowel space. As a result, English infants have a stronger perceptual representation of /u/ in their system compared to the French infants (Molnar & Polka, 2004). It appears that the simpler front vowel system of English compared to French influences perceptual preference for the /u/ vowel by English-learning infants which is in turn reflected in the greater number of /u/ productions by English infants' compared to the French infants.

In summary, the previous research on ambient language influences on babbling has yielded conflicting findings. Some scholars emphasize that changes in babbling are language-general and independent from the influence of the ambient language, whereas other scholars argue that infants learning different languages babble differently as a result of the differences in their linguistic environmental input. One part of the controversy may be due to the differences in the design and measurement tools used by the different studies that examined ambient language effects on babbling: for example, differences in how consonantal inventories are examined, the use of only phonetic transcription, a small sample size, an examination of only one or a few age/language groups, the inclusion of only a few native listeners, the inclusion of bilingual listeners, and findings based on non-statistical analyses. The other part of the controversy may be due to the fact that previous studies (except for Rvachew et al., 2008; Rvachew et al., 2006) searched for evidence for



or against a drift on babbling toward ambient language (Brown's [1958] *babbling drift hypothesis*), instead of searching for evidence for or against the *interactional hypothesis* (de Boysson-Bardies et al., 1989) that claims that infant speech production is a product of an interaction among several variables: infant speech perception, attentional biases, limitations of speech motor control, small vocal tract and speech input. One step towards determining whether the *interactional hypothesis* is valid is to provide evidence that some characteristics of ambient language input influence infant vocal production, as opposed to searching for evidence for a drift in babbling toward an ambient language system (i.e., the infant vowel space matches the input vowel space in a linear fashion).

The present study searched for evidence for the ambient language influences on infant vocal output while taking into consideration the issues related to the controversies raised by earlier findings. The present study expected to find ambient language effects rather than a drift on infant vowel space. Specifically, it was expected that the infant vowel space would be related to the complexity of the vowel system of the ambient language, and that infants learning a simple vowel system would achieve a larger size of the vowel space compared to infants learning a complex vowel system. In addition, the investigation included a comparison of infants exposed to Arabic language at home and raised in an Arabic-speaking city to monolingual infants exposed to the English language at home and to monolingual infants exposed to the French language at home. Compared to many earlier studies, this cross-sectional design included a larger sample size of 31 Arabic infants, 23 French infants, and 20 English infants; and a larger age range from 10 to 18 months. Since English and French are considered to be Indo-European languages with complex vowel systems, and Arabic is a Semitic language with a simple vowel

system (Lewis, 2009), the possibility of finding the prelinguistic influences of ambient language before 12 months of age is maximized. Particularly, early ambient language influences were expected in the size and the expansion of the vowel space: as a result of learning a simple vowel system, Arabic infants were expected to show larger and earlier expansion of their vowel spaces compared to English and French infants. The analysis focused on infant vowel production because, as discussed previously, the language-specific perception of vowels develops as early as 6 months (Kuhl et al., 1992); because accurate vowel production is achieved before consonants (Stoel-Gammon & Pollock, 2008); and because Arabic has a simple vowel system compared to English and French vowel systems, which enables the examination of the relationship between input complexity and changes in the infant vowel space. Furthermore, the vowels were analyzed acoustically, which allowed for better tracking of the developmental and crosslinguistic changes that may not be in evidence when using only auditory-based measurements (e.g., phonetic transcription, perceptual judgment). In addition, an infraphonological coding was used to locate the vowels with measurable formant frequencies. The coding and acoustic analysis was performed by a person with advanced phonetic experience. In addition, the consonantal repertoires of the Arabic, English, and French infants were described and compared. Although many consonants require complex oromotor coordination to produce compared to vowels, language-specific perception for consonants is developed from the ages of 8 to 10 months (Werker & Tees, 2002). This pattern of language-specific perception and the differences of the consonantal inventories of Arabic, English and French may yield early crosslinguistic differences relating to the infant phonetic repertoire. However, it was assumed that such early effects

of ambient language might not be recovered with phonetic transcription analyses, and that crosslinguistic differences in consonantal repertoires might not be identified.

### **1.3 Purpose of the study**

The purpose of the present study was to trace the development of babbling in a large sample of Arabic infants from 10 to 18 months of age. The current investigation contributed knowledge to this end by describing the developmental changes in the vowel spaces and consonantal repertoires of 10- to 18-month-old infants learning Arabic. The Arabic language has received little attention in the literature, and a population of this age range has not been investigated previously using both an acoustic analysis of the vowel space and a phonetic transcription of consonantal inventories with a large sample size of 31 Arabic-learning infants. Such a description enables an examination of a controversial issue in the field—the linguistic environmental influences on babbling—by comparing Arabic developmental data to data from two well-investigated language groups: English and French. By using a cross-sectional design, speech samples were collected from 31 Arabic-learning infants from 10-18 months of age. Babbling data for the Arabic infants were provided through an acoustic analysis of vowels produced in one of the following infraphonological categories or protophones: a fully resonant vowel, marginal babbling, or canonical babbling (these categories are described further in the “Methods” section). In addition, babbling data were provided through phonetic transcription analyses of the consonants produced in canonical babbling and marginal babbling. Furthermore, two crosslinguistic investigations were conducted by using a quantitative causal-comparative research design (Gay, Mills, & Airasian, 2012). The primary investigation involved

comparing Arabic data on the development of the vowel space to the existent dataset of Canadian monolingual English and French infants (Rvachew et al., 2006) to confirm and extend the existent knowledge about the developmental changes in the vowel space (e.g., the early predominance of central vowels and the gradual expansion of the vowel space). In addition, since Arabic, English, and French are phonetically and phonologically distinct languages, the presence of language-specific variations in babbling (in this case, vowel space) among Arabic infants compared to English and French infants contributed to a new line of evidence that illuminates the role of ambient language input during the prelinguistic period and provides evidence to support the *interactional hypothesis* (de Boysson-Bardies et al., 1989). Particularly, the current study expected to find language-specific changes on the center, corners, and size of the vowel spaces among the three language groups. More importantly, this study expected to find that the infant vowel space is related to the complexity of the vowel system of ambient language. That is, it was expected that Arabic infants learning a simple vowel system would show a larger size of the vowel space compared to French and English infants learning more complex vowel systems. The secondary investigation involved crosslinguistic comparisons of the consonantal repertoires of Arabic infants to those of English infants and French infants between the ages of 10-18 months. Given the paucity of descriptive research on Arabic speech development during infancy, the purpose of this investigation was to provide developmental data on the consonantal repertoires from a large scale sample of Arabic infants exposed to Saudi dialect. To my knowledge, only one published study exists that describes the consonantal repertoires of four infants aged 6, 10, 12 and 17 months exposed to Egyptian Arabic (Omar, 1973). In addition, this present investigation on the

consonantal repertoires of Arabic, English, and French infants was to show that early crosslinguistic differences in the vowel space were not accompanied by crosslinguistic differences in consonantal repertoires when analyzed by auditory-based measurements such as phonetic transcription.

The following four questions were addressed in two studies:

### **Study One**

- (1) What are the developmental changes in the acoustic characteristics of the vowel space of Arabic infants between the ages of 10-18 months?
- (2) Do crosslinguistic differences exist in the acoustic characteristics of the vowel space of Arabic, English, and French infants between the ages of 10-18 months? Particularly, do crosslinguistic differences exist in the center, corners, and size of the vowel spaces among the three language groups from 10-18 months?

### **Study Two**

- (3) What are the developmental changes in the consonantal repertoires of Arabic infants between the ages of 10-18 months?
- (4) Do crosslinguistic differences exist in the consonantal repertoires of Arabic, English, and French infants between the ages of 10-18 months?

## 2. ARABIC

The purpose of this chapter is to describe the Arabic phonology system. Section 2.1 presents a brief background of the Arabic language. Section 2.2 describes the phonology of Arabic including the sound systems, syllable structure, stress patterns, and intonation. Then, this chapter ends with a figure that presents a summary of vowels, consonants and stress assignment of English, Arabic, and French.

### 2.1 Background on the Arabic language

Arabic is one of a group of languages known as the Semitic languages, which include, for example, Hebrew and Ethiopian. It is the first language for more than 250 million people, and the official language for 23 countries in the Middle East and North Africa. Arabic also is spoken as second language in many countries where Islam is a majority or minority religion, since the holy Islamic book the *Quran* is written in the Arabic language, and it is read in Arabic even though it has been translated into many other languages. The Arabic language is characterized by a small vowel but large consonant inventory (3 vowels and 28 consonants). Arabic also is considered to be a morphology-rich language. Usually, the lexical forms are derived from the three-consonant root with patterns of vowel changes encoding grammatical functions. For example, from a word with the “**KaTaB**” pattern, which is Arabic for “writing,” different words can be formulated such as “**KiTaaB**” meaning “book,” “**KaaTiB**” meaning “writer,” and “**MaKTuuB**” meaning “letter” or “written.” Also, the nouns and adjectives are inflected to indicate the gender (masculine or feminine) and number (singular, dual, or plural). Arabic is written from right to left by using 28 Arabic alphabetical letters

representing consonants. Vowels are represented by diacritics placed below or above the consonant letters, but they are not used frequently. Arabic script is considered to be a “shallow” orthography, with a direct mapping from graphemes to phonemes (Carter, 2001; Holes, 2004; Watson, 2002).

The Arabic language has several varieties. Classical Arabic is the old formal Arabic, the formal language of the pre-Islamic and early Islamic periods in which the *Quran* and old texts (e.g., poetry and records of historical events) were written. Classical Arabic has not undergone any major changes, and so its uses are limited to the *Quran* and some old or new poetry. Modern Standard Arabic is the modern version of classical Arabic. Usually, it is learned through formal education, and is used for formal communication (e.g., professional meetings, TV news, newspapers). It is not often used for informal conversations between people speaking the same Arabic dialect, but, with dialectal modifications, it may be used to facilitate communication among speakers of different dialects. Colloquial Arabic is the informal version of spoken Arabic, and it is not commonly written except in Internet social network sites. It is the mother tongue of native Arabic speakers. Many varieties of colloquial Arabic exist, and they are distributed geographically across countries (e.g., Egyptian, Syrian, Iraqi) and even within the same country (e.g., Najdi Saudi Arabic, Hijazi Saudi Arabic, bedouin Palestinian Arabic, fellāhi Palestinian Arabic). The dialectal differences increase with increased geographical distance between the dialect groups. For example, the Najdi Saudi dialect (spoken in central Saudi Arabia) and the Qatari dialect are geographically close, and the speakers of these two dialects are able to understand each other without any difficulty. However, the Moroccan dialects and the Najdi Saudi dialect are geographically distant, and the

speakers of these dialects may have difficulty understanding each other, and may need to rely on modern Arabic to communicate (Holes, 2004).

## **2.2 Arabic phonology**

Arabic has many distinct phonological characteristics, but some aspects of its phonology are common across other languages. This section discusses Arabic phonology in comparison to English phonology, which is a well-investigated system. The discussion includes the similarities and differences in the sound systems, syllable structure, stress patterns, and intonation of Arabic and English. The focus is on Modern Standard Arabic with some general notes on Saudi and some other Arabic dialects because many Arabic dialects share many characteristics with the phonology of Modern Arabic and because research is lacking on many Arabic dialects, including Saudi dialects.

### **2.2.1 Vowel system**

Arabic has a small vocalic inventory compared to English (Ladefoged, 2004; Thelwall & Sa'adeddin, 1999). For example, Modern Arabic has three vowels (i.e., /i/, /u/, /a/) that are phonemically contrasted with their long counterparts (i.e., /i:/, /u:/, and /a:/), whereas English has a rich vowel space that consists of 12 vowels (i.e., /i/, /ɪ/, /e/, /ɛ/, /æ/, /a/, /ɔ/, /o/, /ʊ/, /ʉ/, /ʌ/, /ɜ/). Figure 2.1 shows English and Arabic vowels according to their articulatory tongue height and advancement. In addition, Arabic has two diphthongs (i.e., /aw/ and /ay/) that can be reduced to monophthongs to be /e:/ for



/aɪ/ and /oɪ/ for /aw/ (depending on the dialect), whereas English has five diphthongs (/aɪ/, /aʊ/, /ɔɪ/, /eɪ/, and /oʊ/).

The vowel system of some Arabic dialects is more complex, since it includes additional vowels (Holes, 2004), such as /ɛ/, /ə/, /æ/, and /ʊ/ in Egyptian dialects (Omar, 1973) and /e/, /o/, /æ/, and /ə/ in Palestinian dialects (Shahin, 1996). However, Saudi dialects are described to be similar to Modern Arabic. For example, Holes (1990) described the vowels used in the Gulf dialects (from Kuwait to Oman including eastern parts of Saudi Arabia) and analyzed the same vowels sets as Modern Arabic. Similarly, Ingham (2008) describing Najdi Saudi dialect and Abu-Mansour (2007) describing Meccan (Hijazi) Saudi dialect reported the same vowel sets as Modern Arabic which include the following: /i/, /u/, /a/, /iː/, /uː/, and /aː/.

When the vowel space representing tongue height and advancement is described, Arabic appears as a triangular vowel chart, and English as a quadrilateral vowel chart. Unlike the attention paid to English vowels, few studies have investigated the Arabic vowel space. However, Al-Ani's (1970) investigation provides a detailed acoustic description of the Arabic vowels produced by adult Iraqi males. By comparing the F1 and F2 values of Arabic vowels to English vowels (obtained from the study of Hillenbrand, Getty, Clark, & Wheeler, 1995), it appears that the high-front and high-back corner vowels in Arabic and English (i.e., /i/ and /u/) have similar general acoustic features. The vowel /i/ is a close tense, unrounded, high-front vowel characterized by a low F1 and a high F2. The vowel /u/ is a close tense, rounded, high-back vowel characterized by low

F1 and F2. The vowel /a/, which is not commonly produced in English, is an open lax, unrounded, low-center vowel characterized by high F1 and low F2.

### 2.2.2 Consonant System

Arabic has a rich consonantal inventory with 28 consonants. Figure 2.1 shows the consonants of Modern Arabic and English with respect to their manner of articulation, place of articulation, and voicing. The two languages share 17 consonants (/m/, /b/, /f/, /t/, /d/, /n/, /θ/, /ð/, /s/, /z/, /ʃ/, /dʒ/, /j/, /k/, /l/, /h/, and /w/). The English phonemes that do not occur in Arabic are the following: (1) the voiceless bilabial stop (/p/), which is considered to be an allophone of /b/; (2) the voiceless post-alveolar affricate (/tʃ/); (3) the voiced labiodental fricative (/v/); (4) the voiced velar stop (/g/) (this sound occurs in many Arabic dialects but not in Modern Arabic); (5) the voiced post-alveolar fricative (/ʒ/); and (6) the nasal velar (/ŋ/).

Likewise, the Arabic consonantal inventory has some distinct phonemes that do not occur in English, which are the following: (1) /x/ (the voiceless velar fricative accompanied by an uvular trill); (2) /ɣ/ (the voiced velar fricative); (3) /q/ (the voiced uvular stop); (4) /ħ/ (the voiceless pharyngeal fricative); (5) /ʕ/ (the voiced pharyngeal fricative); (6) /ʔ/ (the voiced glottal stop); (7) and /r/ (the voiced alveolar tap-trill). In addition, Arabic has another distinct group of phonemes known as pharyngealized or emphatic consonants (/tˤ/, /dˤ/, /sˤ/, /ðˤ/) that involve a secondary articulatory gesture

(i.e., tongue-root retraction), which distinguishes them from their plain counterparts, a feature similar to English phonemes (/t/, /d/, /s/, /ð/) (Thelwall & Sa'adeddin, 1999). In other words, these emphatic consonants are produced in a manner similar to that used to produce their plain counterparts, but with a simultaneous retraction of the tongue root (Ladefoged & Maddieson, 1996, chap. 10). In addition, /l<sup>ɛ</sup>/ is the counterpart of the plain /l/, which means that the total number of consonants in Arabic is 29 instead of 28; however, /l<sup>ɛ</sup>/ only occurs in one word “/ʔal<sup>ɛ</sup>ah/”, which is the name of “God,” and in its variations such as “/wʔal<sup>ɛ</sup>ah/”(Al-Ani, 1970).

Arabic consonants can be geminated, so the duration of the consonant has phonemic value. *Gemination* means the prolongation of the same consonant or the lengthening of the closure gap of a stop (Al-Ani, 1970). For example, /da ras/ (meaning “studied”) is a disyllabic word with a CV CVC form, whereas /dar ras/ (meaning “taught”) has a CVC CVC form in which the first /r/ is the coda of the first syllable, and the second /r/ is the onset of the second syllable. English does not have gemination, but consonant doubling occurs across the morpheme boundary, for example, in the word “night-time” (Huthaily, 2003).

The consonant system of Arabic dialects is similar to Modern Arabic with slight modifications that include the loss of some phonemic distinctions or a restructuring of the consonant system (Holes, 2004). For example, the Cairo dialect does not have the interdental fricatives (/θ/ and /ð/), which are replaced with /s/ or /t/ for /θ/, and /z/ or /d/ for

/ð/, depending on the lexical uses. The /q/ (the voiced uvular stop) tends to be fronted to be produced as /g/ (the voiced velar stop) in, for example, the Arabian Peninsula dialects; whereas in, for example, Cairo, Lebanese, and Syrian dialects, /q/ tends to be backed to be produced as /ʔ/ (glottal stop). The emphatic /dˤ/ is usually replaced with /ðˤ/ in many dialects, for example, the Arabian Peninsula dialects.

### **2.2.3 Syllable Structure**

The structure of a monosyllabic word is formed by onset (a consonant or consonant cluster) and rime, which consists of a vowel with/without a coda (final consonant or consonant cluster). Unlike English, Modern Arabic and Arabic dialects do not allow for more than two consonant sequences in a cluster. In addition, consonant clustering does not occur in the onset in Modern Arabic and many Arabic dialects (Farwanah, 2007; Mitchell, 1993, Chap.2). Hence, the syllable structure in Arabic can be as follows: CV, CVV, CVC, CVVC, and CVCC. The first three types are the most common among Arabic dialects, and are permissible in Modern Arabic. The CV, CVV, CVC, and CVVC occurs in word-initial, word-medial, and word-final positions, whereas CVCC occurs only in the word-final position (Al-Ani, 1970). Like English, Arabic has monosyllabic, bisyllabic, or multisyllabic word forms.

### **2.2.4 Stress and Intonation**

Both English and Arabic are considered to be “stress-timed” languages. Both languages denote stress to a syllable in a word, and the stress plays an important role in

the phonological system. However, Arabic has a set of simple and predictable rules for stress assignment in a word. Modern Arabic and all Arabic dialects may follow some similar general rules, but some dialects differ with respect to other rules; however, they are predictable most of the time. For example, multisyllabic words of CV patterns place a primary stress on the first syllable and weak stress on the remaining syllables (a summary of these rules can be found in Abdo, 1969; Al-Ani, 1970). Contrary to Arabic, English follows a set of rules, but these are not completely predictable, and so they are complex to some extent. For example, the stress assignment of a word may differ depending on its grammatical class (i.e., whether it is a noun or verb), for example, “an export” and “to export” (Ladefoged, 2004).

In addition, both Arabic and English are intonational languages (Chahal, 2007). Pitch is used differently (i.e., falling, rising, or plateau F<sub>0</sub> contour) to express the different communicative intents of an utterance (e.g., declarative, question, and command). For example, falling melody is used by both languages for a WH-question; in English, low fall is considered as hostile, and high fall as energetic; in Arabic, low fall is a regular question, and high fall represents disbelief or a high degree of interest. In a yes/no question, a rising tone is used in both languages (English uses different variable tones for different yes/no questions, such as the tag question).

Next, Figure 2.1 presents a summary of the differences and similarities in the vowel spaces, the consonant inventories and the stress assignment of English, Arabic, and French (Fougeron & Smith, 1999; Ladefoged, 1999; Thelwall & Sa'adeddin, 1999).

Vowels of English, Arabic, and French based on the articulatory tongue position								
<-- Tongue Advancement -->								
English			Arabic			French		
i		u	i i:		u u:	i y		u
ɪ		ʊ						
e	ɜː	o				e	ə	o
ɛ						ɛ ɛ̃ ø		
	ʌ	ɔ						ɔ ɔ̃
æ		a		a a:		œ œ̃	a	ɑ̃
Diphthongs: aɪ, aʊ, ɔɪ, eɪ & ou			Diphthongs: aw & ay			Diphthongs: —		

&lt;-- Tongue Height --&gt;

The characteristics of English, Arabic, and French consonants										
		Bilabials	Labio-dentals	Dentals	Alveolars	Palatals	Velars	Uvular	Pharyngeal	Glottals
Stops	En	p b		t d			k g			
	Fr	p b		t d			k g			ʔ
	Ar	b		t d tˤ dˤ			k	q		ʔ
Fricatives	En		f v	θ ð	s z	ʃ ʒ				h
	Fr		f v		s z	ʃ ʒ		ʁ		h
	Ar		f	θ ð	s sˤ, z ðˤ	ʃ	x ɣ		ħ ʕ	h
Affricates	En					tʃ dʒ				
	Fr									
	Ar					dʒ				
Nasals	En	m			n		ŋ			
	Fr	m			n	ɲ	ŋ			
	Ar	m			n					
Rhotic	En				ɹ					
Trill	Ar				r					
Lateral	En Ar Fr				l					
Glides	En Ar	w				j				
	Fr	w				j ɥ				

Stress pattern of English, Arabic, and French	
<b>English</b>	Stress-timed language; partially predictable rules for stress assignment
<b>Arabic</b>	Stress-timed language; simple and predictable rules for stress assignment
<b>French</b>	Syllable-timed language; generally stress assignment for phrase-final syllable

Figure 2.1. Summary of consonants, vowels and stress of English, Arabic and French

### 3. METHOD

This chapter discusses the research methodology employed in the investigations of the development of Arabic babbling, and of the crosslinguistic differences in babbling of Arabic, English, and French infants. Two studies were conducted: 1) investigations of infants' vowel space and 2) investigations of infants' consonantal repertoires. The overall aim of these two studies was to trace the development of Arabic speech during infancy and to describe it in relationship to the speech development of English infants and French infants so to examine the linguistic environmental influences on babbling. First, the common subsections of a method section (participants, apparatus, and procedures) are described, and then the procedures specific to each of the two Studies.

#### 3.1 Participants

A total of 31 Arabic-learning infants ranging in age from 10-18 months (281-591 days) participated in the present study. The parents of the participants were native speakers of Arabic. All the infants were recruited from the Saudi Arabian community in Riyadh, Saudi Arabia and Montreal, Canada through informal advertisements distributed among interested parents. Twenty-nine infants were born and raised in Riyadh, Saudi Arabia while two infants were born and raised in Montreal, Canada. All the infants were exposed to only the Arabic language at home, but one infant went to an Arabic-French daycare for less than two months in Montreal. In addition, all the infants were exposed to the Najdi Saudi dialect spoken in central Saudi Arabia. Ten infants were females and 21 infants were males. Initially, 40 Arabic-learning infants were recruited; however, 9 participants were excluded from the study. One participant was excluded after the

recording because he was recovering from a cold episode, which affected the quality of his recorded speech sample. Two participants were fussy and cried during the session. One participant had a flat tympanogram suggestive of otitis media (she was referred to a medical checkup and was diagnosed with *otitis media*). Another participant had a flat tympanogram and a family history of hearing loss. Four participants did not provide enough utterances for analysis (less than 50 codable utterances). The infant development inventory (IDI) was administered to ensure that all the infants had a normal general development (Ireton, 2005). All the 31 participants passed the IDI screening assessment. According to parental reports, all the infants were born full-term and had an unremarkable birth and medical history. The recruited 31 participants had no history of frequent ear infections or hearing impairment, and were healthy during the recording day. Twenty-eight participants met the canonical syllable ratio (CSR) criterion of 0.2 canonical syllables to utterances to be at least at the canonical babbling stage as suggested by Oller (Oller & Eilers, 1988). Three participants did not meet the ratio of canonical syllable to utterances but were included because they provided sufficient babbling. In addition, the study obtained ethics approval from the Institutional Ethics Board (Faculty of Medicine, McGill University, Montreal, Canada) and from the Research Ethics Committee (College of Medicine, King Saud University, Riyadh, Saudi Arabia). The parents were not paid for their infants' participation in the study.

The sample size of the Arabic sample ( $N = 31$ ; age range: 281-591 days) was comparatively equal to the Canadian English sample ( $N = 20$ ; age range: 303-553 days), and the Canadian French sample ( $N = 23$ ; age range: 311-566 days) in Rvachew et al.'s (2006) study that formed the comparison group for the present study. The inclusion



criteria for the English and French participants recruited in the Rvachew et al.'s (2006) study was similar to the inclusion criteria for the Arabic participants. Thorough details about the English and French participants can be found in Rvachew et al. (2006).

### 3.2 Apparatus

Each Arabic infant's parent completed a short language and hearing questionnaire and consent forms. Speech samples were recorded by Pro Tools LE 7.1 (Digidesign, Inc.) and a Sennheiser lapel microphone. Then, the speech samples were acoustically analyzed using Time-Frequency Representation (TFR) software (Avaaz Innovations Inc., 1999). A tympanometer with a 226-Hz probe tone was used to assess the infants' middle ear function (GSI 33 GSI Tymptstar).

### 3.3 Procedure

**Audiological evaluation.** The included 31 participants did not have any sensorineural hearing loss or frequent otitis media. A parental history was obtained to ensure that each infant had a normal hearing level and a normal middle ear (ME) function. In addition, ME function was objectively screened on the recording day by administering the tympanometer with the 226-Hz probe tone. The pass/fail criteria was adopted from the screening criteria suggested by the American Speech-Language-Hearing Association (1997). To pass the criteria, the static admittance ( $Y_{tm}$ ) must be more than 0.2 mmho or the tympanogram width (TW) less than 235 daPa (i.e., a normal tympanogram). If the  $Y_{tm}$  is less than 0.2 mmho or the TW is more than 235 daPa (i.e., a flat tympanogram), the infant failed the screening, since these results are suggestive of

OM. Ear canal volume was also considered for the passing criteria, since large ear canal volume is associated with the presence of tympanic perforation causing mild conductive hearing loss. Infants with an ear canal volume above  $1.0 \text{ cm}^3$ , which also is associated with a flat tympanogram (i.e., abnormal), failed the screening. Twenty-two infants out of the included 31 passed the screening criteria for ME function. Seven infants were not assessed due to their refusal or unavailability of tympanometry on the recording day. Two infants had a flat tympanogram. These infants were included in the study because their parents reported that their infants did not have recurrent ear infections; and it is not uncommon for infants to experience at least one incident of ear infection during the first year of life, with peak incidents occurring between 6-18 months of age (Klein, 2000). Therefore, a decision was made to include the two infants who did not pass the screening for ME function but provided a speech sample of good recording quality and met the CSR of more than 0.2.

**Speech sample recording and collection.** Pro Tools LE 7.1 and a Sennheiser lapel microphone were used to record the speech samples. The Pro Tools recording parameters were set to the sample rate of 44.1 KHz, bit depth of 16 bit, and fader gain of +12 dB. The minimum sample rate in Pro Tools is 44.1 KHz, which needed to be downsampled to 22050 Hz. Sampling rates of 22050 Hz and a quantization of at least 12-bit conversation is recommended for digitizing a speech signal for acoustic analysis (Kent & Read, 2002). Since Pro Tools has built-in amplification, an external amplifier was not used. The built-in amplification was monitored during the recording to avoid too little or too much amplification. The speech samples were collected at the infant's family home while the parent, caregiver, or familiar relative played with the infant (most recordings

were done with the parent). Infants wore a custom-made vest to which the lapel microphone was clipped. For accurate phonetic coding and analysis, the parent was asked to speak as he/she usually spoke to the infant, and to stop talking when the infant vocalized to avoid overlapping. The sole recording session for each infant took approximately 30 to 60 minutes. Although 50 utterances is sufficient for determining a CSR (Rvachew, Creighton, Feldman, & Sauve, 2002), the collected utterances ranged from 95 to 306 with a mean of 180 and a total of 5591 utterances . Utterances were identified as “a noncry, nonreflexive, nonvegetative sound produced within a single breath unit and with a unifying pitch contour” (Rvachew et al., 2002, p. 25).

**Phonetic transcription and infraphonological utterance coding.** The coding and acoustic analysis was performed by the author who is a native Arabic speaker with phonetic experience and speaks English as second language. Each utterance was transcribed according the IPA and then coded according to the infraphonological categories suggested by Oller (1980, 1986, 2000); Rvachew et al. (2002) provide a summary of these infraphonological categories in tabular format, which includes a definition for most of the categories (see table 3.1). In addition, the phonation pattern of each syllable was coded according to Rvachew et al.’s (2002) phonation codes, which are based on Oller (1980, 1986, 2000). Only the syllables with normal phonation were included in the analysis (see table 3.2).

Table 3.1

*Descriptions of Infraphonological Codes*

Category	Description
Canonical syllable	This syllable has at least one consonant (margin), fully resonant vowel (nucleus), smooth formant frequency transition (change) from the margin to the nucleus with a duration between 25 and 120 ms, normal phonation in the nucleus [no evidence of F0 shift, aphonia, biphonation, falsetto, or harmonic doubling], syllable duration between 100 and 500 ms, intensity range equal or more than 30 dB, and intensity differences between margins and nuclei of at least 10 dB.
Marginal syllable	This syllable has consonant-vowel or vowel-consonant syllable shape and meets all criteria for a canonical syllable except for durational requirements (formant transition and/or syllable duration).
Fully resonant vowel	This syllable has vowel-like syllable shape and measurable F1 and F2 with adequate resonance energy above 1200 Hz.
Quasiresonant vowel	This syllable has nasalized vowel with resonance energy below 1200 Hz.
Squeals	This utterance has high pitch ranges above 1000 Hz (falsetto), and it breaks at some points with some vocal cord tension.
Growling	This utterance has low pitch range below 100 Hz and creaky voice.
Raspberries	This utterance has bilabial or labiolingual trills with no vowel.
Whispers	This utterance is perceived as aphonic with invisible harmonics in the spectrogram.
Falsetto	This utterance has high pitch register.
Yells	This utterance has high intensity vocalization.
Ingressive-egressive sequence	This utterance has alternating ingressive-egressive vocalization.

Table 3.2

*Examples of Phonation Codes*

Category	Description
Normal phonation	Nucleus is produced with a periodic energy source with a parallel and evenly spaced series of harmonics in the narrowband FFT spectrogram.
Biphonations	Nucleus is produced with a periodic energy source but with a sudden appearance of non-parallel harmonics in the narrowband FFT spectrogram overlapping with original harmonics, making the syllable perceived as a harsh.
Harmonic doubling	Nucleus is produced with a periodic energy source but with a sudden appearance of parallel harmonics in the narrowband FFT spectrogram overlapping with original harmonics, making the syllable perceived as a harsh.
Aphonic	Nucleus is produced with aperiodic waveform, making the syllable perceived as a whisper.
Quasiresonant	Nucleus is produced with resonance energy below 1200 Hz, making the syllable perceived as nasalized.

### 3.4 Method of Study One: Crosslinguistic investigation of infants' vowel space

The aim of this study was to describe the acoustic vowel space of 10- to 18-month-old Arabic infants in relation to the acoustic vowel space of 10- to 18-month-old English infants and French infants. Particularly, the analyses were based on the mean of the first formant frequency ( $M_{F1}$ ); the mean of the second formant frequency ( $M_{F2}$ ); standard deviation of the first formant frequency ( $SD_{F1}$ ); the standard deviation of the second formant frequency ( $SD_{F2}$ ); the most grave, most acute, most compact, and most diffuse vowels; and vowel space area. The results highlight the direction of changes in the center of the vowel space and the expansion of the vowel space as a function of age and language background. The acoustic analyses and data analyses used in Study One are described in the following paragraphs.

**Participants.** The participants included 31 Arabic infants, 23 French infants, and 20 English infants from 281-591 days of age. Details about participants' characteristics and selection criteria are discussed in subsection 3.1.

**Acoustic analysis.** Pro Tools was used to splice the speech samples into utterances. These utterances were down-sampled from 44.1 KHz to 22.050 KHz for acoustic analysis, according to Kent and Read's (2002) suggestions. These utterances were processed by Praat software (Boersma & Weenink, 2011) to be compatible for TFR processing. Then, the TFR was used to produce spectrograms of the utterances from which the F1 and F2 could be identified. Vowels that were produced in canonical syllable (CS), Marginal syllable (MS), and Fully resonant vowel (FRV) were considered for the

analysis. The criteria for identifying these infraphonological categories were presented in the preceding section. All vowels produced in CS were included, and only the vowels produced in MS and FRV with normal phonation and measurable F1 and F2 were included. The syllable selection criteria for the Arabic sample were loosened to include all CS and many MS and FRV compared to the selection criteria for the English and French samples used in Rvachew et al.'s (2006) study, which only included CS and FRV. The decision to include MS was made because some Arabic participants produced few CS but many MS with normal phonation. As described previously in Table 3.1, MS with normal phonation meets all CS criteria except for the durational requirements (formant transition and/or syllable duration); such a violation of CS criteria should not impact the calculation of the F1 and F2 of MS with normal phonation. The total number of analyzed segments was 6773; 35.7% were CS, 27.4% were MS and 36.9% were FRV. The details about English and French samples are found in Rvachew et al. (2006). The selected vowels from CS, MS and FRV had measurable F1 and F2 that were used to derive the infant vowel space. F1 and F2 were measured from 20 ms of steady state in the middle of the vowel. The TFR software was set to produce a linear predictive coding (LPC) autocorrelation analysis with a processing parameter of 256-point window size, 50%-overlap, 98 %-preemphasis, hanning window type, and a 12-model order (the model order was adjusted depending on the measurability of the formant frequencies). The narrowband short-time fast fourier with similar LPC parameters—except for a 512-point window size—was used to recheck the locations of F1 and F2 in the spectrogram for some vowels for which the location of F1 and F2 was not certain based on LPC analysis. The data for F1 and F2 for English children provided by Hillenbrand et al. (1995) were

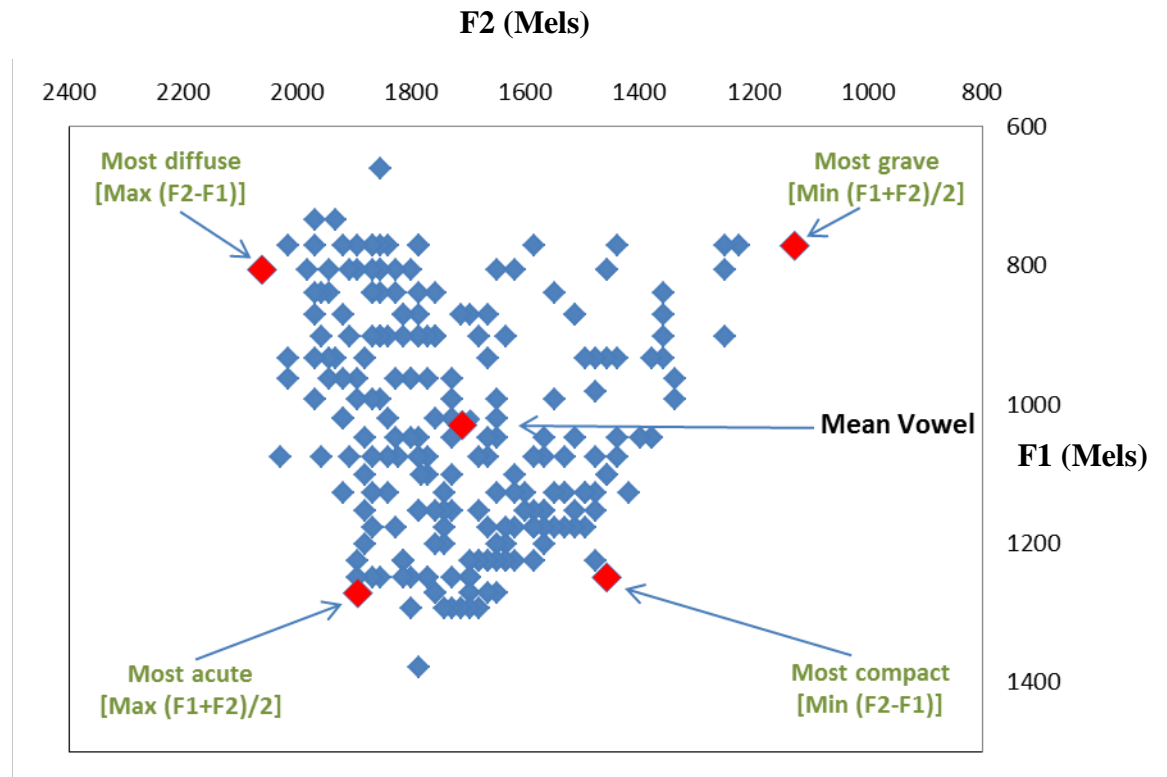
used to aid in the identification of F1 and F2 locations in the spectrogram. Syllable duration was measured using a waveform display. The measurement started from the point just before the amplitude rose from zero energy for a consonant in CV or CVC or a vowel in VC to the point where the amplitude fell to zero energy. The F2 transition was measured using a spectrogram generated by the LPC analysis. The procedure for the acoustic analysis was adopted from the study by Rvachew et al. (2006).

**Data Analysis.** The Arabic developmental data on vowel space were compared to the developmental data of English and French infants. Accordingly, the data were analyzed to examine the effects of age and language background on the vowels produced by the Arabic, English, and French infants. For all infants, F1 and F2 values were determined for each vowel meeting the selection criteria and were converted from Hertz to Mel scale to account for the non-linearity frequency perception by human ear and to represent the vowel space in a form similar to human ear perception (Reetz & Jongman, 2009) using the formula:  $F_{\text{Mel}} = (1127.010481) \ln [ 1 + ( F_{\text{Hertz}} / 700) ]$  (Stevens, Volkmann, & Newman, 1937). In addition, the compact-diffuse ( $F2 - F1$ ) and grave-acute ( $(F2+F1)/2$ ) values were calculated for each vowel. The compact-diffuse and grave-acute dimensions provided linear distribution of the infant produced vowels and that allowed for a description of the infant vowel space in relation to the adult vowel space, making possible the tracking of the developmental changes of the infant vowel space (Kuhl & Meltzoff, 1996). The following were identified: the most acute vowel—the greatest value of the grave-acute dimension [i.e., maximum value of  $(F1+F2)/2$ ], the most grave—the least value of the grave-acute dimension [i.e., minimum value of  $(F1+F2)/2$ ], the most diffuse—the greatest value of the compact-diffuse dimension [i.e., maximum value ( $F2 -$

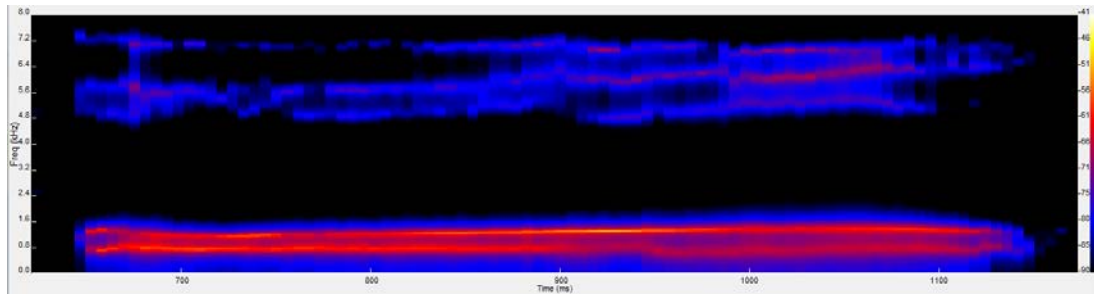


F1)], the most compact—the least value of the compact-diffuse dimension [i.e., minimum value (F2 - F1)]. In addition, the  $M_{F1}$ ,  $M_{F2}$ ,  $SD_{F1}$ , and  $SD_{F2}$  were calculated for each infant's sample. Figure 3.1 provides an example of an infant vowel space with the four corner vowels and the mean vowel identified, and Figure 3.2 presents the spectrograms of these four corner vowels.

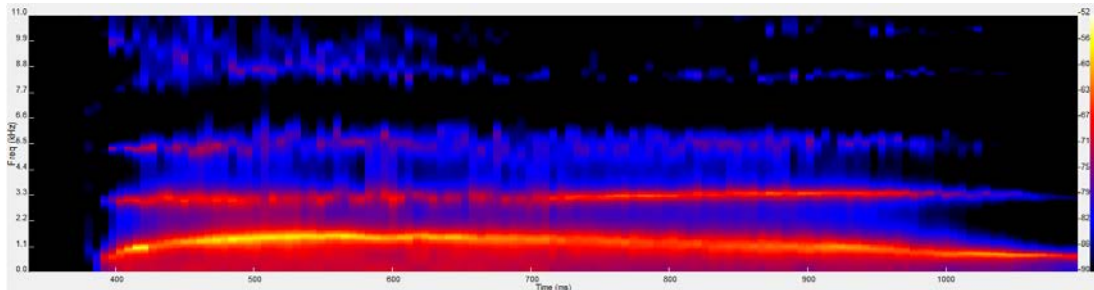
In addition, the triangular area size of the infant vowel space was identified for each infant. This area represented a working vowel space that was calculated for each infant by using Liu, Tsao, and Kuhl's (2005) equation: triangle vowel space = absolute value of  $( [ F1i * ( F2a - F2u ) + F1a * ( F2u - F2i ) + F1u * ( F2i - F2a ) ] / 2 )$  where /i/, /a/ and /u/ were replaced with extreme diffuse, compact, and grave vowels, respectively. The extreme vowels were identified based on the extreme value, and then confirmed by a visual inspection of each infant vowel space; the selected vowels were found on the corner of the vowel space and were not outliers. The visual inspection and data sorting approaches showed agreement on identifying the diffuse and grave vowels except for a few cases (three Arabic vowels, one English vowel, and five French vowels). For each of these few cases, the second or the third vowel with extreme values was chosen. In addition, the selected compact vowels were found with minimum 10% value of (F2-F1) except for nine vowels were found below 20% and six vowels were found below 27%. The selection of compact vowels with some flexibility was done to ensure the representation of more than 90% of infant produced vowels. Figure 3.3 shows an example of an infant vowel space with the three corners of most diffuse, grave and compact vowels (the compact vowel has a minimum 16% value of [F2-F1]).



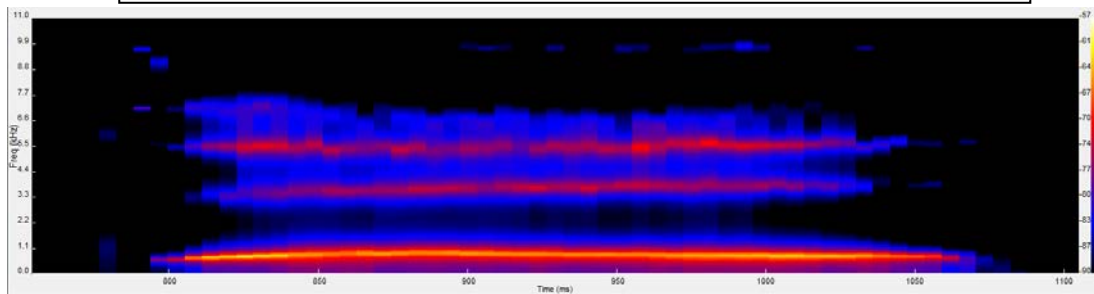
*Figure 3.1.* An example of the vowel space of a 14-month-old Arabic infant showing the mean vowel and the most diffuse, grave, acute, and compact vowels.



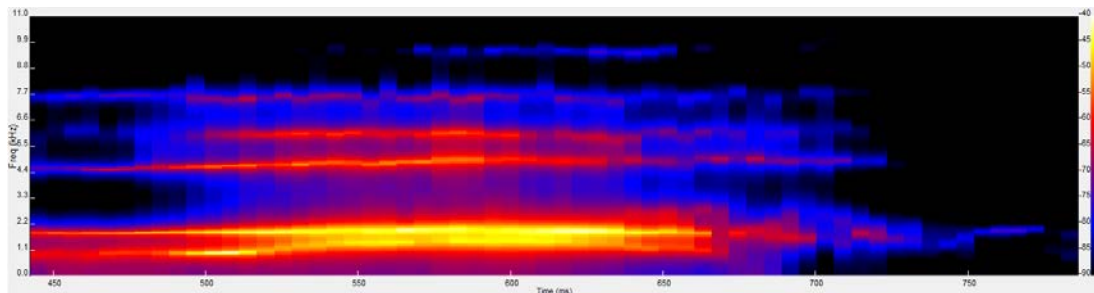
Extreme grave vowel (/u/-like): F1= 689 Hz, F2= 1205 Hz



Extreme acute vowel (/æ/-like): F1= 1464 Hz, F2= 3057 Hz

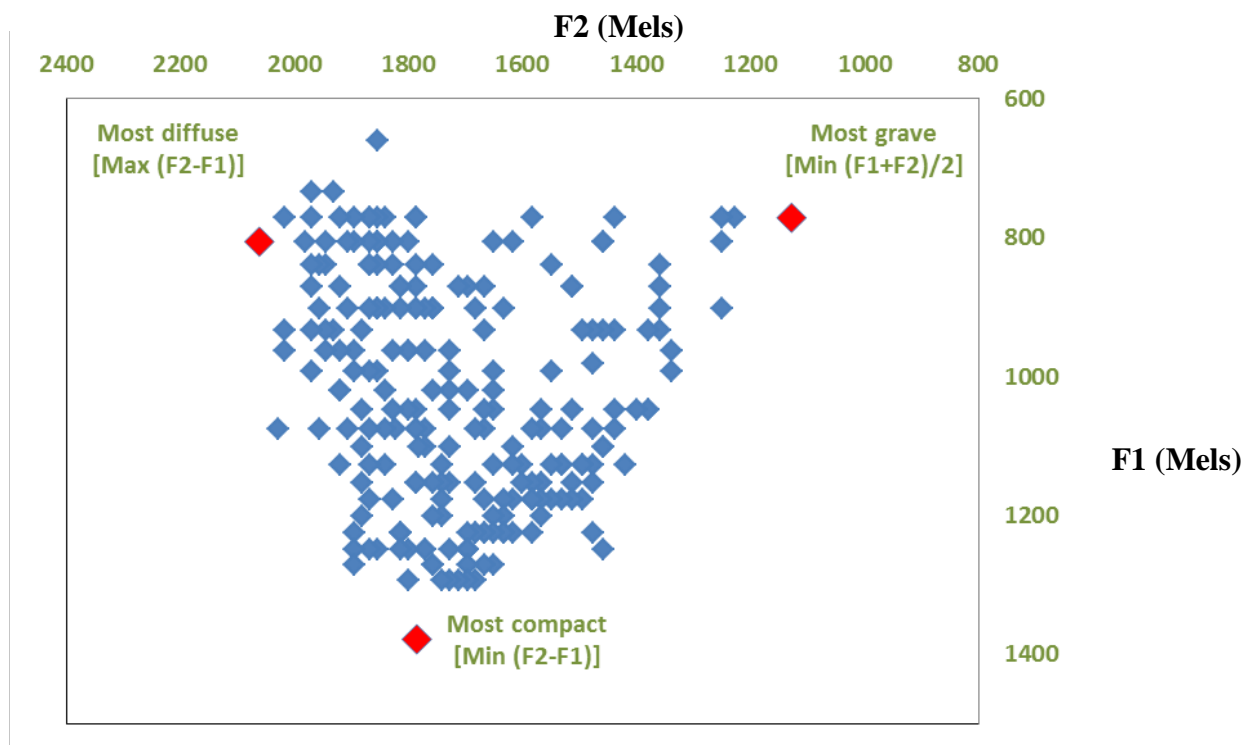


Extreme diffuse vowel (/i/-like): F1= 732 Hz, F2= 3660 Hz



Extreme compact vowel (/a/-like): F1= 1421 Hz, F2= 1851 Hz

*Figure 3.2.* Spectrograms of extreme vowels (corner vowels) of a 14-month-old Arabic infant.



*Figure 3.3.* An example of the triangular vowel space of a 14-month-old Arabic infant showing the three corners of the most diffuse, grave, and compact vowels.

Multiple regression analyses were conducted to examine the effects of age (predictor/ independent variable [IV1]), language (IV2), and an age-language interaction on the following dependent variables (DV): the vowel space parameters  $M_{F1}$ ,  $M_{F2}$ ,  $SD_{F1}$ ,  $SD_{F2}$ , and the features derived from those parameters, specifically most acute, most grave, most diffuse, and most compact values. Two series of the multiple regression analyses were conducted: 1) to examine the crosslinguistic differences between the Arabic and French groups; and 2) to examine the crosslinguistic differences between the Arabic and English groups. In the current study, the English and French groups were not compared directly with each other in the current study because they were compared previously in Rvachew et al. (2006), and because the aim of the current study was to compare the Arabic vowel space to the vowel spaces of two languages of different rhythmic classes (English is a stress-timed language, and French is a syllable-timed language). The IVs were coded using effect coding as suggested by Pedhazur (1997) to examine the interaction among IVs. Subsequently, simple regression analyses were conducted to examine the effect of age on each DV for each language group. The simple regression analyses were also used to construct a regression model to estimate the value of each DV at the youngest age (i.e., 281 days) and the oldest age (i.e., 591 days) for each crosslinguistic comparison using the following model:  $Y_i = b_0 + (b_1 X)$ , where  $Y_i$  is the predicated value of DV,  $b_0$  and  $b_1$  are regression coefficients, and  $X$  is an infant's age in days (281 or 591) for a particular DV (Field, 2009). In addition, a two-way between-subject analysis of variance (ANOVA) was conducted to examine the effects of language (IV1), effects of age (IV2), and the effect of language-age interaction on the triangular area size of the infant vowel space. For ANOVA testing, the infants were organized into

three age groups: 10-12 months (281-352 days), 12-15 months (378-478 days), 15-18 months (486-591 days).

### 3.5 Method of Study Two: Crosslinguistic investigation of infants' consonantal repertoires

The aim of this study was to describe the consonantal repertoire of 10- to 18-month-old Arabic infants in relation to the consonantal repertoire of 10- to 18-month-old English infants and French infants. The results highlight the differences and similarities in the proportions of phonetic categories as a function of language background. The phonetic transcription and infraphonological utterance coding were the same as those used in Study One. The data analyses used in Study Two are described in the following paragraphs.

**Participants.** The same groups of infants recruited for Study One participated in this study, that is, 31 Arabic infants, 20 English infants, and 23 French infants ranging in age from 10-18 months (281-591 days). The Infants were organized into three age groups as follows:

- Group 1: 10-12 (months) from 281-352 (days) or 9.4-11.7 (months)
  - N= 24; Arabic n= 12, English n = 6, and French n = 6
- Group 2: 12-15 (months) From 378-478 (days) or 12.2-15.9 (months)
  - N= 27; Arabic n= 12, English n = 8, and French n = 7
- Group 3: 15-18 (months) From 486-591 (days) or 16.2-19.7 (months)
  - N= 23; Arabic n= 7, English n = 6, and French n = 10

**Data Analysis.** The Arabic developmental data on consonantal repertoire were obtained through crosslinguistic comparisons with English infants and French infants.

Accordingly, data were analyzed to examine the effects of language background on the consonantal repertoires produced by the Arabic, English, and French infants. To describe the phonetic repertoire, only the consonants that were produced in CS and MS with normal phonation were included in the analysis to provide an accurate transcription. The total number of consonants was 6484 (4440 Arabic, 1113 French and 931 English consonants). The consonants were grouped in phonetic categories according to the manner and place of production as is shown in Figure 3.2. The percentage of each phonetic category was calculated for each infant by dividing the number of consonants in each phonetic category by the total numbers of consonants that an infant produced and then multiplied by 100.

Statistical analyses were done on the proportion of each place of articulation category (labials, coronals, dorsal, and glottal) and each manner of articulation category (stops, nasals, fricatives, glides, liquids). Two-way between-subject analysis of variance (ANOVA) was conducted to examine the effects of language (IV1), effects of age (IV2), and the effect of language-age interaction on each DV: percentages of each place category (labials, coronals, dorsal, and glottal) and each manner category (stops, nasals, fricatives, glides, liquids).



		Place			
Manner		Labials	Coronals	Dorsal	Glottal
	Stops	b p	t d	k g q	ʔ
	Nasals	m	n	ɳ	
	Fricatives	f v	θ ð s z ʃ ʒ tʃ tʂ dʒ ʤ	ɣ x ɣ̞	ħ h
	Glides	w	j		
	Liquids		r l		

*Figure 3.2.* Grouping of consonants

## 4. RESULTS

This chapter presents the results of the crosslinguistic investigation on babbling, which described the course of speech development among Arabic infants in relation to English infants and French infants. First, the results of Study One (the investigation of infants' vowel space) are presented, and then the results of study Two (the investigation of infants' consonantal repertoires) are described.

### 4.1 Results of Study One: Crosslinguistic investigation of infants' vowel space

This section presents the results of the crosslinguistic comparisons of the acoustic characteristics of the vowel space of Arabic, English, and French infants between the ages of 10-18 months. The results of this subsection are presented in the following order: results of reliability analyses, results of multiple regressions analyses on the center and corners of vowels space, and results of ANOVA on the size of infant vowel space.

#### 4.1.1 Reliability

The F1 and F2 values of 13% of the random Arabic sample were re-checked by a second individual trained in phonetic and acoustic analysis by using the same acoustic analysis method. The intraclass correlation coefficient was 0.952 for F1 and 0.92 for F2, which suggested an acceptable level of agreement.

#### **4.1.2 The analyses of the center and corners of the vowel space: Arabic versus French**

By using the Enter method, multiple regression analyses were conducted to examine the effects of age, language, and language-age interaction on the parameters (i.e.,  $M_{F1}$ ,  $M_{F2}$ ,  $SD_{F1}$ , and  $SD_{F2}$ ) and features of the vowel space (i.e., extreme values of the acute, grave, diffuse and compact features) of Arabic and French infants from 10-18 months of age. Table 4.1 presents a summary of descriptive statistics for the criterion variables (vowel space parameters and features). Presented next are the results of these multiple analyses and the simple regression analyses that examined the separate effect of age on the criterion variables for each language group. The results of the analyses of the parameters of the vowel space are presented first followed by the results of the analyses of the features of the vowel space.

Table 4.1

*Descriptive statistics (in Mels) for each of the vowel space parameters and features for Arabic, English, and French infants between the ages of 10-18 months*

	English		Arabic		French	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
$M_{F1}$	867.94	106.26	886.61	72.09	837.28	135.72
$SD_{F1}$	168.21	31.19	164.67	32.61	161.93	34.41
$M_{F2}$	1627.07	111.17	1664.46	94.80	1650.51	100.56
$SD_{F2}$	182.28	61.93	177.85	47.12	166.59	53.14
Most Acute	1471.99	73.54	1558.31	53.28	1458.16	88.07
Most Grave	982.98	139.68	928.99	75.90	975.71	112.19
Most Diffuse	1204.67	231.61	1396.11	128.80	1274.31	242.88
Most Compact	374.29	122.05	287.40	90.17	391.03	212.21

### *Parameters of vowel space*

The following sets of analyses are concerned with the parameters of the vowel space (i.e.,  $M_{F1}$ ,  $M_{F2}$ ,  $SD_{F1}$ , and  $SD_{F2}$ ). The results highlight the effects of age, language, and language-age interaction in the F1 and F2 dimensions of the vowel space.

The  $M_{F1}$  is related to changes in tongue height during the production of vowels (i.e., lower values of F1 correspond to greater tongue height in high vowels such as /i/ and /u/ and *vice versa*). The results of the multiple regressions showed that the three predictors—age, language, and language-age interaction—significantly explained 24.5% of the variance in the  $M_{F1}$  value [adjusted  $R^2 = .2$ ,  $F(3, 50) = 5.4$ ,  $p = .003$ ]. A significant language effect ( $B = -151.32$ ,  $t = -2.18$ ,  $p = .034$ ), a significant age effect ( $B = -0.47$ ,  $t = -2.94$ ,  $p = .005$ ), and a significant language-age interaction effect ( $B = 0.39$ ,  $t = 2.47$ ,  $p = 0.017$ ) were found. As is illustrated in Figure 4.1 (top left), as age increased, French infants showed a significant decline in the  $M_{F1}$  value from 979 to 713 [ $B = -0.86$ ;  $SE = 0.3$ ;  $F(1, 21) = 8.02$ ,  $p = .01$ ], whereas Arabic infants showed a smaller non-significant decline from 896 to 873 [ $B = -0.08$ ;  $SE = 0.16$ ;  $F(1, 29) = 0.23$ ,  $p = .64$ ]. Thus, an examination of Figure 4.1 (top left) suggests that French infants lowered their F1 values with age to produce vowels in the adult's /i-/u/ dimension compared to Arabic infants who tended to keep their F1 values relatively stable. The divergence on the  $M_{F1}$  values between the two languages (i.e., lowering on French  $M_{F1}$  compared to constant Arabic  $M_{F1}$ ) is noted around the age of 12 months.

The  $M_{F2}$  value is related to changes in tongue advancement during the production of vowels (i.e., higher values of F2 correspond to greater advancement of the tongue in

front vowels such as /i/ and /æ/ and *vice versa*). The results of the multiple regressions showed that the three predictors—age, language, and language-age interaction—did not significantly explain the variance in the  $M_{F2}$  value [adjusted  $R^2 = -0.03$ ,  $F(3, 50) = 0.43$ ,  $p = .73$ ]. Therefore, as illustrated in Figure 4.1 (top right), the decline in the  $M_{F2}$  value for Arabic infants from 1689 to 1629 [ $B = -0.192$ ;  $SE = 0.204$ ;  $F(1, 29) = 0.88$ ,  $p = .36$ ] and French infants from 1669 to 1635 [ $B = -0.11$ ;  $SE = 0.26$ ;  $F(1, 21) = 0.18$ ,  $p = .68$ ] was not significant. Thus, an examination of Figure 4.1 (top right) suggests that both groups of infants kept their F2 values relatively stable with age.

The  $SD_{F1}$  value is related to changes in the range or value of F1 or tongue height. The results of the multiple regressions showed that the three predictors—age, language, and language-age interaction—did not significantly explain the variance in the  $SD_{F1}$  value [adjusted  $R^2 = 0.02$ ,  $F(3, 50) = 1.31$ ,  $p = .28$ ]. Therefore, as illustrated in Figure 4.1 (bottom left), the changes in the  $SD_{F1}$  value for the Arabic infants from 151 to 185 [ $B = 0.11$ ;  $SE = 0.07$ ;  $F(1, 29) = 2.47$ ,  $p = .13$ ] and the French infants from 179 to 147 [ $B = -0.1$ ;  $SE = 0.09$ ;  $F(1, 21) = 1.39$ ,  $p = .25$ ] were not significant. Thus, these data show that both groups of infants tended to have more stability in the F1 dimension or in the range of vertical tongue positions with age.

The  $SD_{F2}$  value is related to changes in the range or value of F2 or tongue advancement. The results of the multiple regressions showed that the three predictors—age, language, and language-age interaction—significantly explained 14.1% of the variance in the  $SD_{F2}$  value [adjusted  $R^2 = .09$ ,  $F(3, 50) = 2.74$ ,  $p = .053$ ]. Only a significant age effect [ $B = 0.17$ ,  $t = 2.14$ ,  $p = .037$ ] was found. As illustrated in Figure 4.1

(bottom right), as age increased, the Arabic infants showed a significant increase in the  $SD_{F2}$  value from 144 to 228 [ $B = 0.273$ ;  $SE = 0.09$ ;  $F(1, 29) = 9.21$ ,  $p = .005$ ], whereas the French infants showed a non-significant change from 156 to 176 [ $B = 0.07$ ;  $SE = 0.14$ ;  $F(1, 21) = 0.23$ ,  $p = .63$ ]. The divergence on  $SD_{F2}$  value (i.e., increasing Arabic  $SD_{F2}$  value compared to stable French  $SD_{F2}$  value) is noted prior to the age of 12 months. Thus, these data show that Arabic infants have more variability in the F2 dimension with age, which suggests that Arabic infants were expanding their range of horizontal tongue positions during speech (i.e., tongue advancement), compared to French infants who showed more stability.

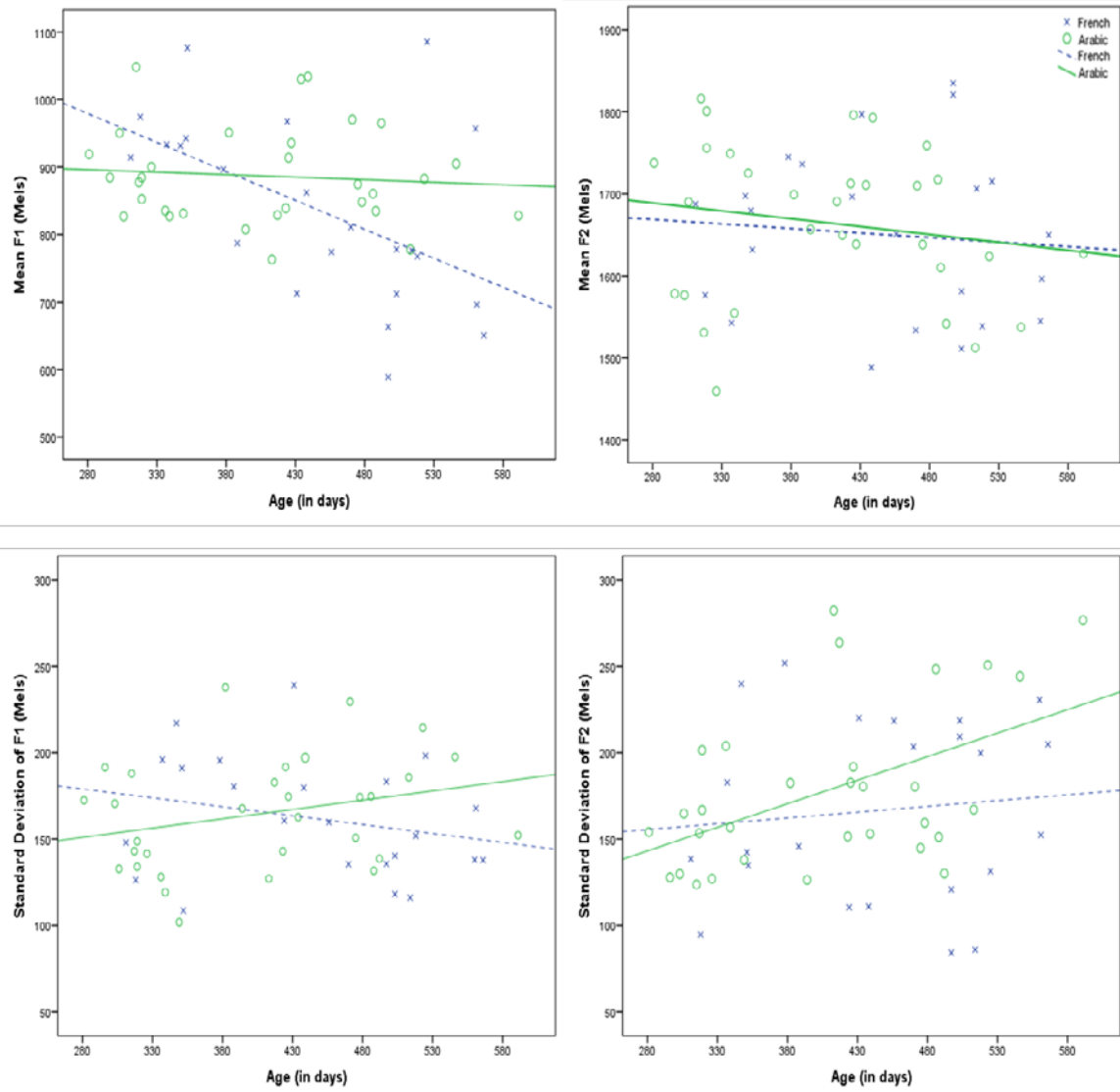


Figure 4.1. The  $M_{F1}$  (top left),  $M_{F2}$  (top right), the  $SD_{F1}$  (bottom left), and  $SD_{F2}$  (bottom right) values plotted for each infant as a function of age (281 through 591 days) and language group (Arabic versus French).



### *Features of the vowel space*

The following sets of analyses are concerned with the features of the vowel space (i.e., extreme values of the acute, grave, diffuse, and compact features). The results highlight the effects of age, language, and language-age interaction on the grave-acute dimension (adult's /u/-/æ/) dimension and the compact-diffuse dimension (adult's /a/-/i/ dimension) of the vowel space.

The acute feature of the vowel space (/æ/-like vowel) represents the greatest value of the grave-acute dimension [i.e., maximum value of  $(F1+F2)/2$ ] that, in the adult talker, is bounded by the low-front vowels characterized by high F1 and F2 values. The results of the multiple regressions showed that the three predictors—age, language, and language-age interaction—significantly explained 48% of the variance in the *acuteness* value [adjusted  $R^2 = .449$ ,  $F(3, 50) = 15.38$ ,  $p < .001$ ]. A significant age effect ( $B = -0.28$ ,  $t = -2.67$ ,  $p = .01$ ), a significant language-age interaction effect ( $B = 0.31$ ,  $t = 2.92$ ,  $p = .005$ ), and a marginal language effect ( $B = -87.85$ ,  $t = -1.89$ ,  $p = .065$ ) were found. As illustrated in Figure 4.2 (bottom left), as age increased, the French infants showed a significant decline in *acuteness* value from 1556 to 1371 [ $B = -0.6$ ;  $SE = 0.19$ ;  $F(1, 21) = 9.65$ ,  $p = .005$ ], whereas the Arabic infants showed relatively no changes in the *acuteness* from 1555 to 1563 across the age range [ $B = 0.03$ ;  $SE = 0.12$ ;  $F(1, 29) = 0.05$ ,  $p = .82$ ]. The divergence on the *acuteness* value between the two languages is noted from the age of 10 months. Thus, these data suggest that French infants shrunk their vowel space from the acute corner into the center with age, whereas the Arabic infants appeared to keep their *acuteness* value stable with age.

The grave corner of the vowel space (/u/-like vowel) represents the least value of the grave-acute dimension [i.e., minimum value of  $(F1+F2)/2$ ] that, in the adult talker, is bounded by the high-back vowels characterized by low F1 and F2 values. The results of the multiple regressions showed that the three predictors—age, language, and language-age interaction—significantly explained 25.2% of the variance in the *graveness* value [adjusted  $R^2=.207$ ,  $F(3, 50) = 5.62$ ,  $p = .002$ ]. Only a significant age effect ( $B = -0.51$ ,  $t = -3.57$ ,  $p = .001$ ) was found. As illustrated in Figure 4.2 (top right), as age increased, the decline in the *graveness* value for the Arabic infants from 986 to 885 [ $B = -0.46$ ;  $SE = 0.14$ ;  $F(1, 29) = 10.23$ ,  $p = .003$ ] and the French infants from 1067 to 895 [ $B = -0.56$ ;  $SE = 0.27$ ;  $F(1, 21) = 4.27$ ,  $p = .051$ ] was significant. Thus, these data suggest that both groups of infants expanded their vowel space into the grave corner, and that Arabic infants appeared to have relatively extreme values of *graveness* at a younger age and throughout the age range compared the French infants

The diffuse feature of the vowel space (/i/-like vowel) represents the greatest value of the compact-diffuse dimension [i.e., maximum value  $(F2 - F1)$ ] that, in the adult talker, is bounded by the high-front vowels characterized by low F1 and high F2 values. The results of the multiple regressions showed that the three predictors—age, language, and language-age interaction—significantly explained 22.8% of the variance in the *diffuseness* value [adjusted  $R^2 = .181$ ,  $F(3, 50) = 4.91$ ,  $p = .005$ ]. A significant age effect [ $B = 0.82$ ,  $t = 2.78$ ,  $p = .008$ ] and a marginal language effect ( $B = 237.28$ ,  $t = 1.85$ ,  $p = .071$ ) were found. As illustrated in Figure 4.2 (top left), as age increased, the French infants showed a significant increase in the *diffuseness* value from 1078 to 1448 [ $B = 1.19$ ;  $SE = 0.58$ ;  $F(1, 21) = 4.19$ ,  $p = .053$ ], whereas the Arabic infants showed a smaller

non-significant increase from 1341 to 1477 across the age range [ $B = 0.438$ ;  $SE = 0.27$ ;  $F(1, 29) = 2.64$ ,  $p = .12$ ]. Thus, an examination of Figure 4.2 (top left) suggests that both groups of infants expanded their vowel space into the diffuse corner with age, although the Arabic infants showed a non-significant increase. In addition, Arabic infants appear to have achieved relatively extreme values of *diffuseness* at a younger age than the French infants.

The compact feature of the vowel space (/a/-like vowel) represents the least value of the compact-diffuse dimension [i.e., minimum value ( $F2 - F1$ )] that, in the adult talker, is bounded by the low-back vowels characterized by high  $F1$  and low  $F2$  values. The results of the multiple regressions showed that the three predictors—age, language, and language-age interaction—significantly explained 17.1% of the variance in the *compactness* value [adjusted  $R^2 = .122$ ,  $F(3, 50) = 3.45$ ,  $p = .023$ ]. However, only a marginal effect of the language-age interaction ( $B = -0.48$ ,  $t = -1.91$ ,  $p = .062$ ) was found. As illustrated in Figure 4.2 (bottom right), as age increased, the Arabic infants showed a significant decline in the *compactness* value from 360 to 182 [ $B = -0.57$ ;  $SE = 0.17$ ;  $F(1, 29) = 11.85$ ,  $p = .002$ ], whereas the French infants showed relatively no changes in the *compactness* value from 326 to 448 across the age range [ $B = 0.39$ ;  $SE = 0.55$ ;  $F(1, 21) = 0.51$ ,  $p = .48$ ]. The divergence on the *compactness* value between the two languages is noted prior the age of 12 months. Thus, these data suggest that Arabic infants expanded their vowel space into the compact corner with age, whereas the French infants appeared to keep their *compactness* value stable with age.

The findings from the regression analyses between Arabic and French groups are summarized in Table 4.2. The crosslinguistic similarities and differences in the pattern of changes (i.e., increase, decrease, or no change) in the vowel space parameters and features as a function of age are highlighted in Table 4.2.

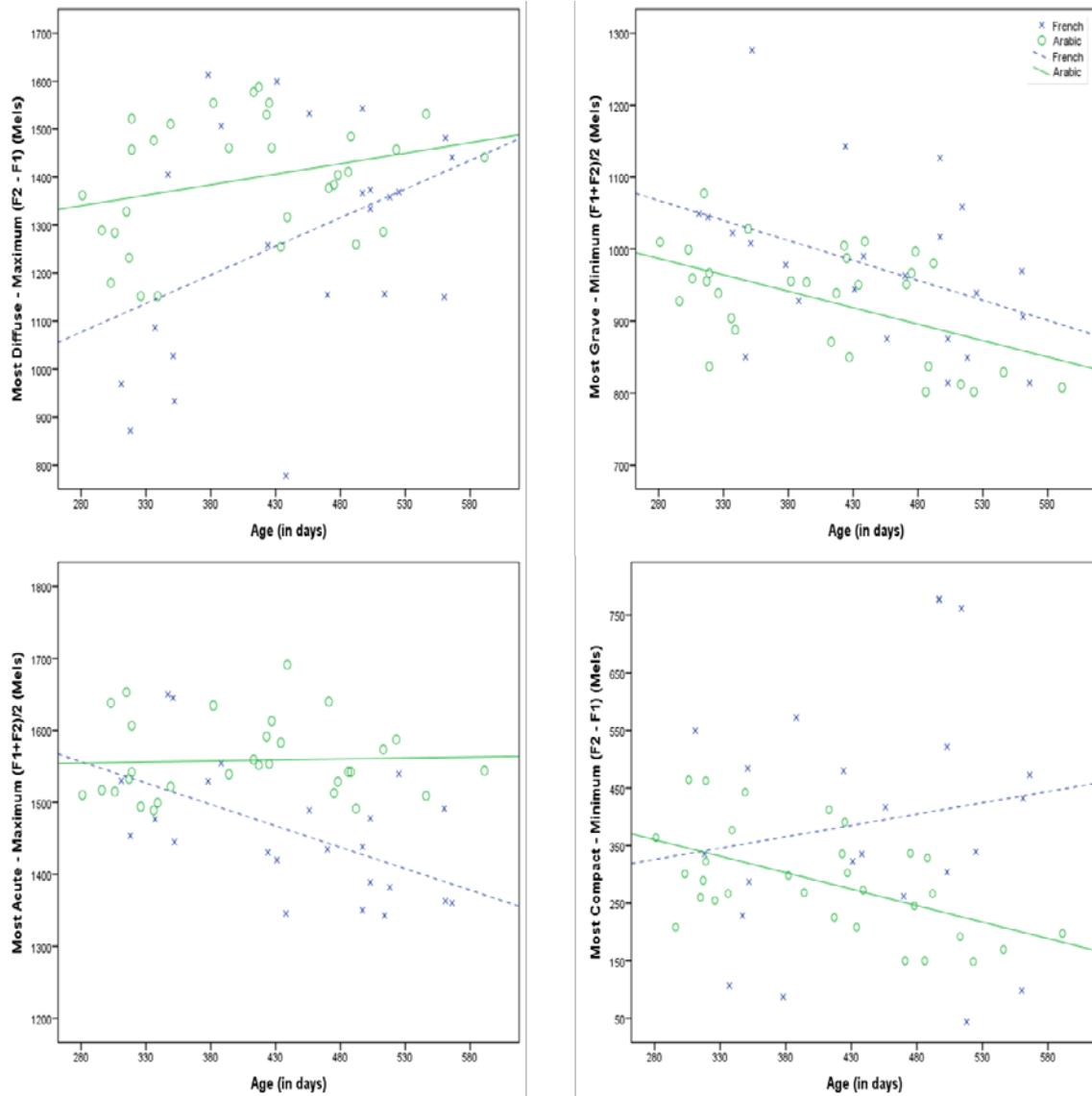


Figure 4.2. The most diffuse (top left), most compact (bottom right), most acute (bottom left), and most grave (top right) vowels plotted for each infant as a function of age (281 through 591 days) and language group (Arabic versus French).

Table 4.2

*Summary of the results from the regression analyses. The asterisk (\*) symbol means a significant effect on the value of the criterion variable (i.e.,  $p < .054$ ). Trend means that a marginal effect occurred in the value of the criterion variable (i.e.,  $p \leq 0.1$  to  $p = .055$ ). The em dash (—) symbol means that no significant changes occurred across the age range.*

		Effect of Predictor Variables			Change Pattern	
		Language	Age	Interaction	Arabic	French
Parameters	$M_{F1}$	*	*	*	—	↓
	$M_{F2}$				—	—
	$SD_{F1}$				—	—
	$SD_{F2}$		*		↑	—
Features	Most Acute	Trend	*	*	—	↓
	Most Compact			Trend	↓	—
	Most Diffuse	Trend	*		—	↑
	Most Grave		*		↓	↓

### 4.1.3 The analyses of the center and corners of the vowel space: Arabic versus English

By using the Enter method, multiple regression analyses were conducted to examine the effects of age, language, and language-age interaction on the parameters (i.e.,  $M_{F1}$ ,  $M_{F2}$ ,  $SD_{F1}$ , and  $SD_{F2}$ ) and features of the vowel space (i.e., extreme values of the acute, grave, diffuse, and compact features) of Arabic and English infants from 10-18 months of age. Table 4.1 presents a summary of the descriptive statistics for the criterion variables (parameters and features of the vowel space). Presented next are the results of these multiple analyses and the simple regression analyses that examined the separate effect of age on the criterion variables for each language group. The results of the analyses of the parameters of the vowel space are presented first followed by the results of the analyses of the features of the vowel space.

#### *Parameters of vowel space*

The following sets of analyses are concerned with the parameters of the vowel space (i.e.,  $M_{F1}$ ,  $M_{F2}$ ,  $SD_{F1}$ , and  $SD_{F2}$ ). The results highlight the effects of age, language, and language-age interaction on the F1 and F2 dimensions of the vowel space

The  $M_{F1}$  is related to changes in the tongue height during the production of vowels. The results of the multiple regressions showed that the three predictors—age, language, and language-age interaction—did not significantly explain the variance in the  $M_{F1}$  value [adjusted  $R^2 = 0.01$ ,  $F(3, 47) = 1.16$ ,  $p = .34$ ]. Therefore, as illustrated in Figure 4.3 (top left), the decline in the  $M_{F1}$  value for Arabic infants from 896 to 873 [ $B = -$

0.08;  $SE = 0.16$ ;  $F(1, 29) = 0.23$ ,  $p = .64$ ] and English infants from 920 to 806 [ $B = -0.37$ ;  $SE = 0.27$ ;  $F(1, 18) = 1.89$ ,  $p = .17$ ] was not significant. Thus, an examination of Figure 4.3 (top left) suggests that both groups of infants kept their F1 values relatively stable with age.

The  $M_{F2}$  value is related to changes in the tongue advancement during the production of vowels. The results of the multiple regressions showed that the three predictors—age, language, and language-age interaction—significantly explained 19.4% of the variance in the  $M_{F2}$  value [adjusted  $R^2 = .142$ ,  $F(3, 47) = 3.77$ ,  $p = .017$ ]. A significant age effect ( $B = -0.45$ ,  $t = -2.84$ ,  $p = .007$ ) and a trend of a language-age interaction effect ( $B = 0.26$ ,  $t = 1.63$ ,  $p = .109$ ) were found. As illustrated in Figure 4.3 (top right), as age increased, English infants showed a significant decline in the  $M_{F2}$  value from 1728 to 1508 [ $B = -0.71$ ;  $SE = 0.24$ ;  $F(1, 18) = 8.64$ ,  $p = .009$ ], whereas Arabic infants showed a smaller non-significant decline from 1689 to 1629 [ $B = -0.19$ ;  $SE = 0.2$ ;  $F(1, 29) = 0.88$ ,  $p = .36$ ]. Thus, an examination of Figure 4.3 (top right) suggests that English infants lowered their F2 values with age to produce more front vowels compared to Arabic infants who tended to keep their F2 values relatively stable. The divergence on the  $M_{F2}$  values between the two languages (i.e., lowering on English  $M_{F2}$  compared to constant Arabic  $M_{F2}$ ) is noted from the age of 12 months.

The  $SD_{F1}$  value is related to changes in the range or value of F1 or tongue height. The results of the multiple regressions showed that the three predictors—age, language, and language-age interaction—did not significantly explain the variance in the  $SD_{F1}$  value [adjusted  $R^2 = 0.001$ ,  $F(3, 47) = 1.01$ ,  $p = .39$ ]. Therefore, as illustrated in Figure



4.3 (bottom left), the increase in the  $SD_{F1}$  value for Arabic infants from 151 to 185 [ $B = 0.11$ ;  $SE = 0.07$ ;  $F(1, 29) = 2.47$ ,  $p = .13$ ] and the decrease in  $SD_{F1}$  value for English infants from 176 to 160 [ $B = -0.05$ ;  $SE = 0.08$ ;  $F(1, 18) = 0.41$ ,  $p = .53$ ] were not significant. Thus, these data show that both groups of infants tended to have more stability in the F1 dimension or in the range of vertical tongue positions with age.

The  $SD_{F2}$  value is related to changes in the range or value of F2 or tongue advancement. The results of the multiple regressions showed that the three predictors—age, language, and language-age interaction—significantly explained 31.4% of the variance in the  $SD_{F2}$  value [adjusted  $R^2 = .27$ ,  $F(3, 47) = 7.16$ ,  $p < .001$ ]. Only a significant age effect [ $B = 0.35$ ,  $t = 4.62$ ,  $p < .001$ ] was found. As illustrated in Figure 4.3 (bottom right), as age increased, the increase in  $SD_{F2}$  value for the Arabic infants from 144 to 228 [ $B = 0.27$ ;  $SE = 0.09$ ;  $F(1, 29) = 9.21$ ,  $p = .005$ ] and the increase for the English infants from 121 to 254 [ $B = 0.43$ ;  $SE = 0.13$ ;  $F(1, 18) = 10.91$ ,  $p = .004$ ] were significant. Thus, these data show that both groups had more variability in the F2 dimension with age, which suggests that these infants were expanding their range of horizontal tongue positions during speech (i.e., tongue advancement).

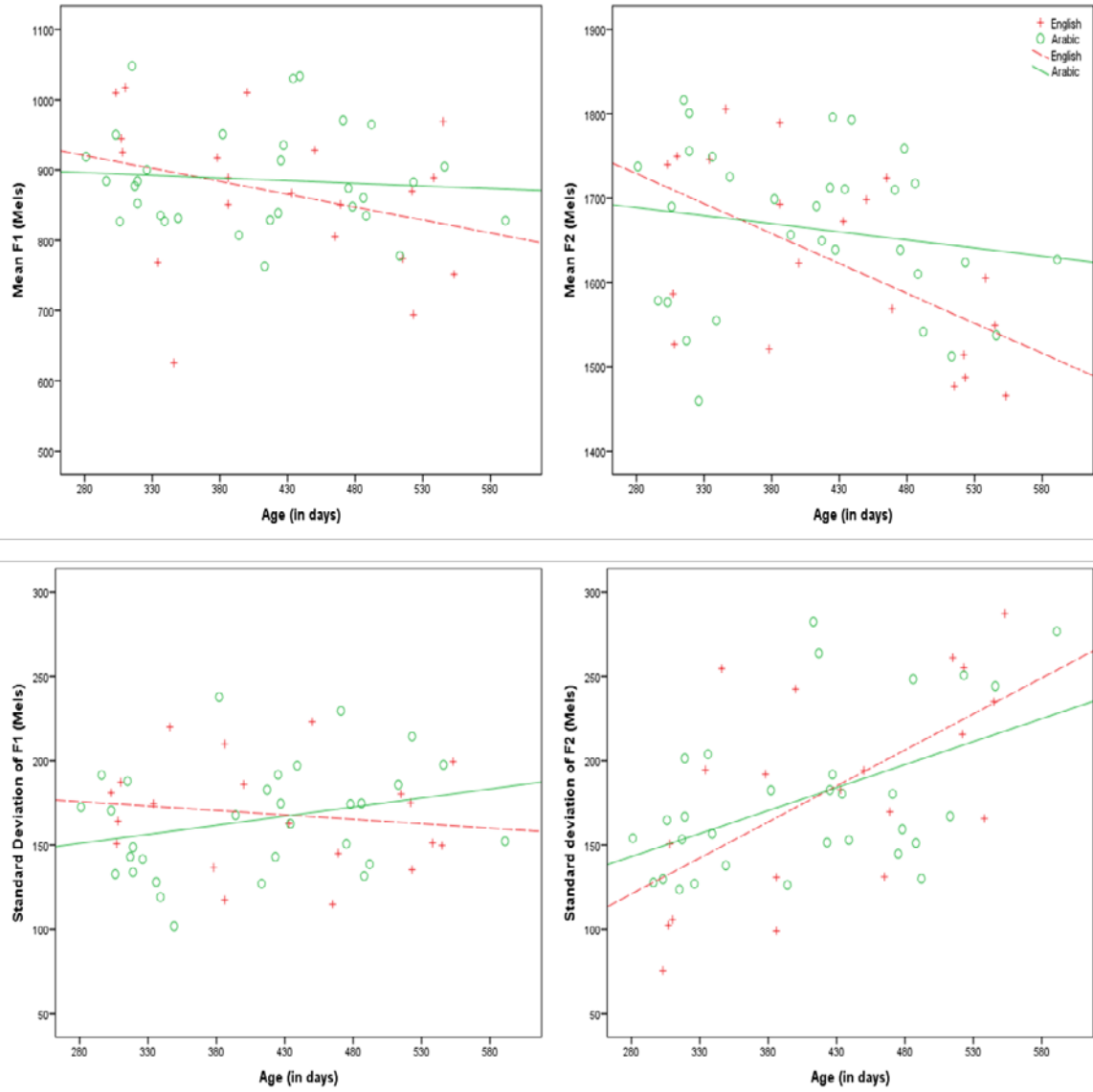


Figure 4.3. The  $M_{F1}$  (top left),  $M_{F2}$  (top right),  $SD_{F1}$  (bottom left), and  $SD_{F2}$  (bottom right) plotted for each infant as a function of age (281 through 591 days) and language group (Arabic versus English).

### *Features of the vowel space*

The following sets of analyses are concerned with the features of the vowel space (i.e., extreme values of the acute, grave, diffuse, and compact features). The results highlight the effects of age, language, and language-age interaction on the grave-acute dimension and the compact-diffuse dimension of the vowel space.

The acute feature of the vowel space (/æ/-like vowel) represents the greatest value of the grave-acute dimension [i.e., maximum value of  $(F1+F2)/2$ ] that, in the adult talker, is bounded by the low-front vowels characterized by high F1 and F2 values. The results of the multiple regressions showed that the three predictors—age, language, and language-age interaction—significantly explained 39.4% of the variance in the *acuteness* value [adjusted  $R^2 = .356$ ,  $F(3, 47) = 10.2$ ,  $p < .001$ ]. However, a marginal effect of language-age interaction ( $B = 0.19$ ,  $t = 1.9$ ,  $p = .063$ ) and a trend of age effect ( $B = -0.16$ ,  $t = 1.64$ ,  $p = .109$ ) were found. As illustrated in Figure 4.4 (bottom left), as age increased, the English infants showed a marginal significant decline in the *acuteness* value from 1523 to 1413 [ $B = -0.36$ ;  $SE = 0.18$ ;  $F(1, 18) = 4.1$ ,  $p = .058$ ], whereas the Arabic infants showed a relatively stable value from 1555 to 1563 across the age range [ $B = 0.03$ ;  $SE = 0.12$ ;  $F(1, 29) = 0.05$ ,  $p = .82$ ]. Thus, these data suggest that both groups of infants had less expansion toward the acute corner of the vowel space.

The grave corner of the vowel space (/u/-like vowel) represents the least value of the grave-acute dimension [i.e., minimum value of  $(F1+F2)/2$ ] that, in the adult talker, is bounded by the high-back vowels characterized by low F1 and F2 values. The results of the multiple regressions showed that the three predictors—age, language, and language-

age interaction—significantly explained 37.7% of the variance in the *graveness* value [adjusted  $R^2 = .338$ ,  $F(3, 47) = 9.49$ ,  $p < .001$ ]. A significant language effect ( $B = -136.47$ ,  $t = -2.18$ ,  $p = .034$ ), a significant age effect ( $B = -0.71$ ,  $t = -4.81$ ,  $p < .001$ ), and a trend effect of language-age interaction ( $B = 0.25$ ,  $t = 1.7$ ,  $p = .096$ ) were found. As illustrated in Figure 4.4 (top right), as age increased, the Arabic infants showed a small but significant decrease in the *graveness* value from 986 to 885 [ $B = -0.46$ ;  $SE = 0.14$ ;  $F(1, 29) = 10.23$ ,  $p = .003$ ], whereas the English infants showed a significant but greater decrease from 1119 to 823 [ $B = -0.96$ ;  $SE = 0.29$ ;  $F(1, 18) = 10.67$ ,  $p = .004$ ]. Thus, these data suggest that both groups of infants expanded their vowel space into the grave corner, and that Arabic infants appeared to have relatively extreme values of *graveness* at a younger age than the English infants.

The diffuse feature of the vowel space (/i/-like vowel) represents the greatest value of the compact-diffuse dimension [i.e., maximum value ( $F2 - F1$ )] that, in the adult talker, is bounded by the high-front vowels characterized by low  $F1$  and high  $F2$  values. The results of the multiple regressions showed that the three predictors—age, language, and language-age interaction—significantly explained 28.9% of the variance in the *diffuseness* value [adjusted  $R^2 = .243$ ,  $F(3, 47) = 6.36$ ,  $p = .001$ ]. Only a significant age effect [ $B = 0.58$ ,  $t = 2.01$ ,  $p = .05$ ] was found. As illustrated in Figure 4.4 (top left), as age increased, the increase in the *diffuseness* value for Arabic infants from 1341 to 1477 across the age range [ $B = 0.438$ ;  $SE = 0.27$ ;  $F(1, 29) = 2.64$ ,  $p = .12$ ] and for English infants from 1101 to 1326 [ $B = 0.73$ ;  $SE = 0.59$ ;  $F(1, 18) = 1.53$ ,  $p = .23$ ] were not significant. An examination of Figure 4.4 (top left) suggests that both groups of infants expanded their vowel space into the diffuse corner with age—however, this expansion

was not statistically significant. In addition, Arabic infants appeared to have achieved relatively extreme values of *diffuseness* at a younger age and throughout the age range compared to the English infants.

The compact feature of the vowel space (/a/-like vowel) represents the least value of the compact-diffuse dimension [i.e., minimum value ( $F_2 - F_1$ )] that, in the adult talker, is bounded by the low-back vowels characterized by high  $F_1$  and low  $F_2$  values. The results of the multiple regressions showed that the three predictors—age, language, and language-age interaction—significantly explained 31.9% of the variance in the *compactness* value [adjusted  $R^2 = .276$ ,  $F(3, 47) = 7.36$ ,  $p < .001$ ]. Only a significant age effect ( $B = -0.53$ ,  $t = -3.32$ ,  $p = .002$ ) was found. As illustrated in Figure 4.4 (bottom right), as age increased, the Arabic infants showed a significant decline in the *compactness* value from 360 to 182 [ $B = -0.57$ ;  $SE = 0.17$ ;  $F(1, 29) = 11.85$ ,  $p = .002$ ], whereas the English infants showed a non-significant decline from 443 to 294 across the age range [ $B = -0.48$ ;  $SE = 0.3$ ;  $F(1, 18) = 2.55$ ,  $p = .13$ ]. Arabic infants had lower *compactness* value from the age of 10 months and continued to achieve relatively extreme values throughout the age range. Thus, these data suggest that Arabic infants expanded their vowel space into the compact corner with age, whereas the English infants appeared to keep their *compactness* value stable with age.

The findings from the regression analyses between Arabic and English groups are summarized in Table 4.3. The crosslinguistic similarities and differences in the pattern of changes (i.e., increase, decrease, or no change) in the vowel space parameters and features as a function of age are highlighted in Table 4.3.

Last, table 4.4 presents a summary of the findings of the regression analyses on the crosslinguistic comparisons of infants' vowel space of the three language groups. It highlights the common changes and the language-specific variations in infants' vowel spaces.

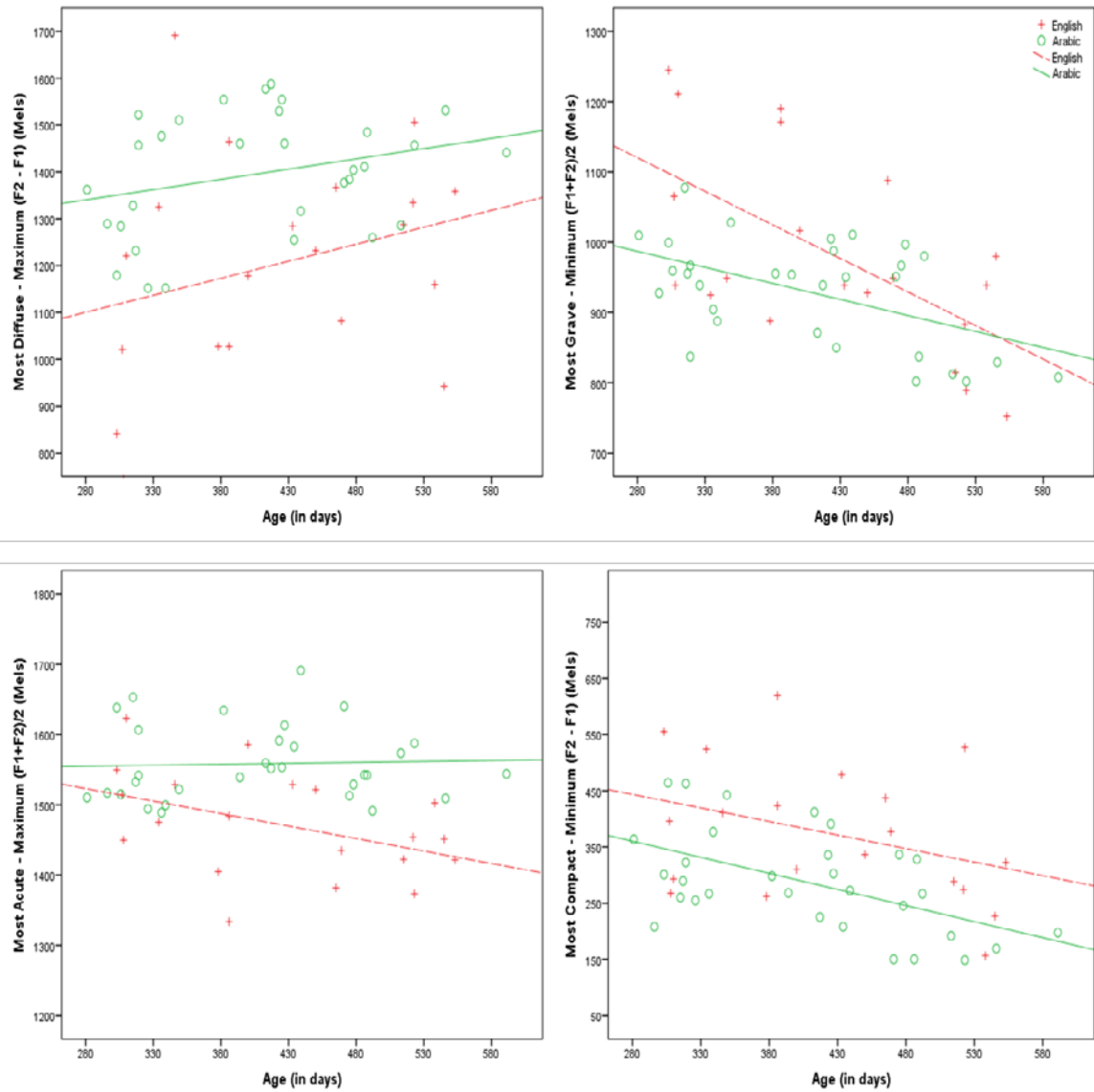


Figure 4.4. The most diffuse (top left), most compact (bottom right), most acute (bottom left), and most grave (top right) vowels plotted for each infant as a function of age (281 through 591 days) and language group (Arabic versus English).

Table 4.3

*Summary of the results from the regression analyses. The asterisk (\*) symbol means a significant effect on the value of the criterion variable (i.e.,  $p < .054$ ). Trend means that a marginal effect has occurred in the value of the criterion variable (i.e.,  $p \leq 0.1$  to  $p = .055$ ). The em dash (—) symbol means that no significant changes have occurred across the age range.*

		Effect of Predictor Variables			Change Pattern	
		Language	Age	Interaction	Arabic	English
Parameters	$M_{F1}$				—	—
	$M_{F2}$		*	Trend	—	↓
	$SD_{F1}$				—	—
	$SD_{F2}$		*		↑	↑
Features	Most Acute		Trend	Trend	—	Trend (↓)
	Most Compact		*		↓	—
	Most Diffuse		*		—	—
	Most Grave	*	*	Trend	↓	↓



Table 4.4

*Summary of the results from the regression analyses of the first study on the crosslinguistic comparisons of infants' vowel space of the three language groups. The em dash (—) symbol means that no significant changes have occurred across the age range.*

		Change Pattern		
		French	Arabic	English
Parameters	$M_{F1}$	↓	—	—
	$M_{F2}$	—	—	↓
	$SD_{F1}$	—	—	—
	$SD_{F2}$	—	↑	↑
Features	Most Acute	↓	—	—
	Less extreme values			
	Most Compact	—	↓	—
	More extreme values			
	Most Diffuse	↑	—	—
	More extreme values			
	Most Grave	↓	↓	↓
More extreme values				

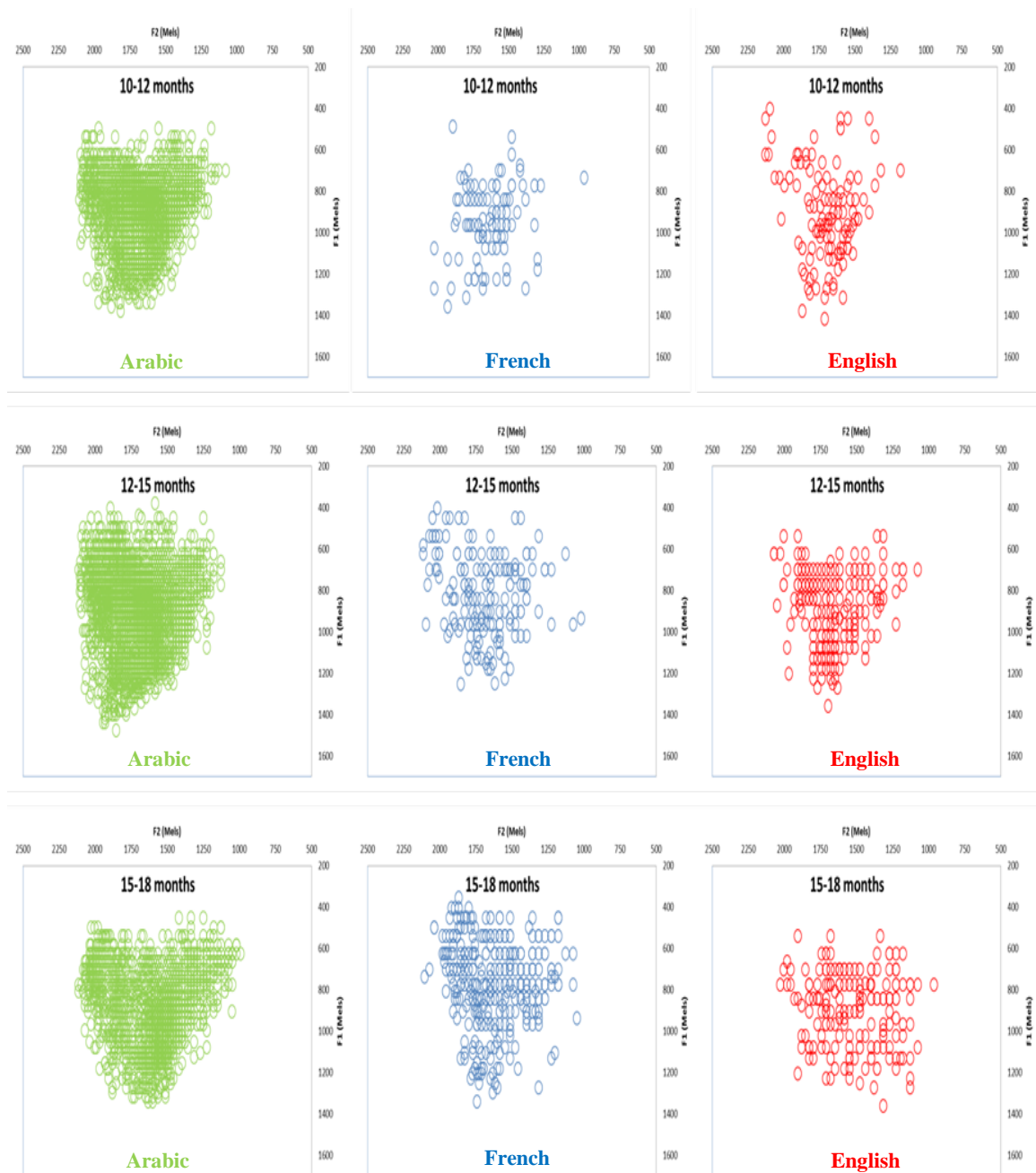
#### 4.1.4 The analyses of the size of the vowel space: Arabic versus French versus English

Two-way between-subject ANOVAs were conducted to examine the effects of age, language, and language-age interaction on the size of the vowel space of Arabic, English, and French infants from 10-18 months of age. Table 4.5 presents a summary of the descriptive statistics for each language group. The analyses showed significant effects of language,  $F(2, 65) = 28.51, p < .001, \eta^2 = .47$ , and significant effects of age,  $F(2, 65) = 11.79, p < .001, \eta^2 = .27$ . One-way ANOVAs conducted for each age interval show significant effects of language group for the infants aged 10-12 months,  $F(2,21) = 18.5, p < .001, \eta^2 = .64$ , and for the infants aged 12-15 months,  $F(2,24) = 14.05, p = .011, \eta^2 = .54$ , and for the infants aged 15-18 months,  $F(2,20) = 5.23, p = .015, \eta^2 = .34$ . The Scheffe's test showed significant differences in the means of the size of the vowel space between 10-12-month-old Arabic and English ( $p < .001$ ), 10-12-month-old Arabic and French ( $p < .001$ ), 12-15-month-old Arabic and English infants ( $p = .003$ ), 12-15-month-old Arabic and French ( $p < .001$ ), and 15-18-month-old Arabic and French ( $p = .015$ ). Considering the significant ANOVA outcomes, the examination of Figure 4.5 and 4.6, and the achievement of extreme *graveness*, *diffuseness*, and *compactness* values by Arabic infants—as found by multiple regression analyses—Arabic infants have achieved a larger size of the vowel space area at a younger age compared to the other two language groups.

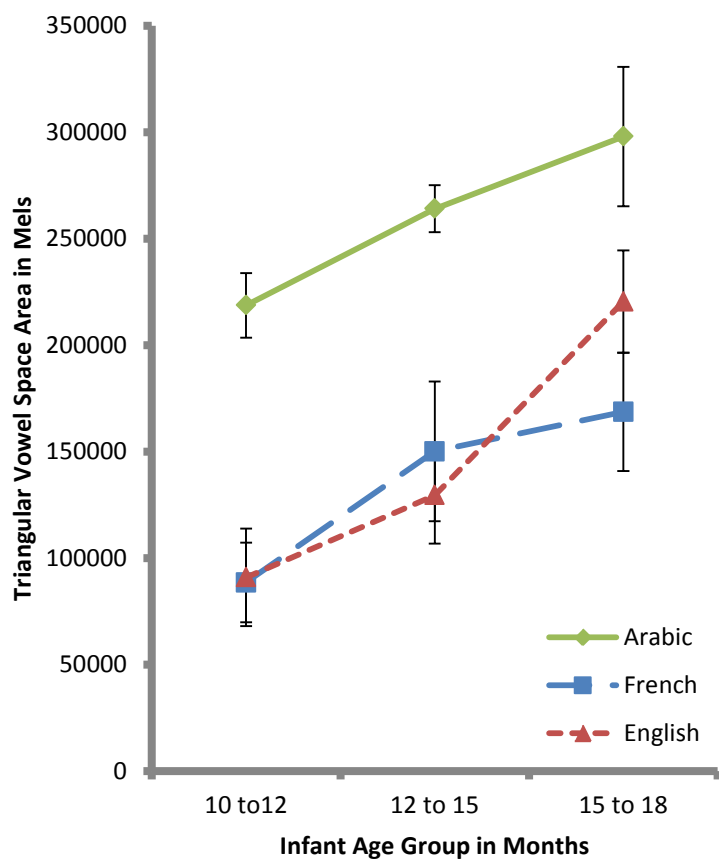
Table 4.5

*Descriptive statistics (in Mels) of the size of the vowel space for Arabic, English, and French infants between the ages of 10-18 months*

	Age (in months)	<i>M</i>	<i>SD</i>	<i>n</i>
Arabic	10 to 12	218734.49	52573.74	12
	12 to 15	264084.94	38201.29	12
	15 to 18	297987.29	86619.23	7
	Total	254185.30	63550.90	31
French	10 to 12	88629.07	45810.03	6
	12 to 15	150186.27	86722.99	7
	15 to 18	168714.05	87795.93	10
	Total	142183.43	82485.99	23
English	10 to 12	90956.89	56092.71	6
	12 to 15	129575.44	64368.04	8
	15 to 18	220496.51	58814.89	6
	Total	145266.20	78019.85	20



*Figure 4.5.* The F1-F2 vowel spaces (in Mels) of Arabic, French, and English infants at 10-12 months (281-352 days), 12-15 months (378-478 days), and 15-18 months (486-591 days).



*Figure 4.6.* Mean triangular area of infant vowel space (in Mels, with standard error bars) plotted as a function of age and language group.

## **4.2 Results of Study Two: Crosslinguistic investigation of infants' consonantal repertoires**

This section presents the results of the crosslinguistic comparisons of the consonantal repertoires of Arabic, English, and French infants between the ages of 10-18 months in three subsections: results of reliability analyses, results for place of articulation and results for manner of articulation.

### **4.2.1 Reliability**

Consonant reliability was checked using a point-to-point agreement approach. A second transcriber with phonetic experience re-transcribed a random 20% of the English and French samples. The inter-rater reliability of the 20 % of the English and French sample was 78% for place of articulation and 80% for the manner of articulation. In addition, the consonant reliability for the Arabic sample was checked by using same approach; also, the same transcriber re-transcribed a random 20% of the Arabic sample one year after the initial transcription. The intra-rater reliability of 20% of the Arabic sample was 87% for the place of articulation and 84% for the manner of articulation. These figures suggest that the reported agreement levels are acceptable given that the infant transcription is a challenging task.

### **4.2.2 Results for place of articulation**

Two-way between-subject ANOVAs were conducted to examine the effects of age, language, and language-age interaction on the proportions of coronals, dorsals,

glottals, and labials of Arabic, English, and French infants from 10-18 months of age.

Table 4.6 presents a summary of descriptive statistics for each place of articulation category, and Figure 4.7 shows the proportion of each place category as plotted by language groups across the three age groups (10-12, 12-15, and 15-18 months). The analyses show that no language, age, and language-age interaction effects occurred in the proportions of production in each place category across language groups, except for an age effect on dorsals, a trend of language effect on coronals, and a trend of language-age interaction effect on labials. The results of the ANOVAs are presented in Table 4.7. The Scheffe's test showed only a significant difference in the means of the dorsal proportion between 10-12 months and 15-18 months, which suggests that an increase occurred in the production of dorsals across the three language groups ( $p = .026$ ).

Table 4.6

*The descriptive statistics in percentages for each place category of Arabic, English, and French infants from 10-18 months of age*

<b>10-12 months</b>						
	Arabic		English		French	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Coronal	64.25	21.87	50.94	29.82	54.54	29.79
Dorsal	7.15	8.53	6.70	4.02	0.88	2.15
Glottal	8.41	8.39	8.87	8.90	6.65	7.18
Labial	20.19	16.72	33.49	29.30	37.93	29.75

<b>12-15 months</b>						
	Arabic		English		French	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Coronal	59.54	17.28	60.78	23.75	47.82	18.09
Dorsal	5.07	6.14	8.73	7.06	9.83	9.65
Glottal	6.00	4.90	6.98	8.93	3.68	2.05
Labial	29.39	18.06	23.51	18.39	38.67	20.35

<b>15-18 months</b>						
	Arabic		English		French	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Coronal	46.74	23.34	48.15	14.51	71.96	19.73
Dorsal	15.38	17.96	13.27	5.79	9.69	6.25
Glottal	7.07	4.15	7.25	8.78	2.20	3.09
Labial	30.81	12.83	31.33	15.78	16.15	14.03



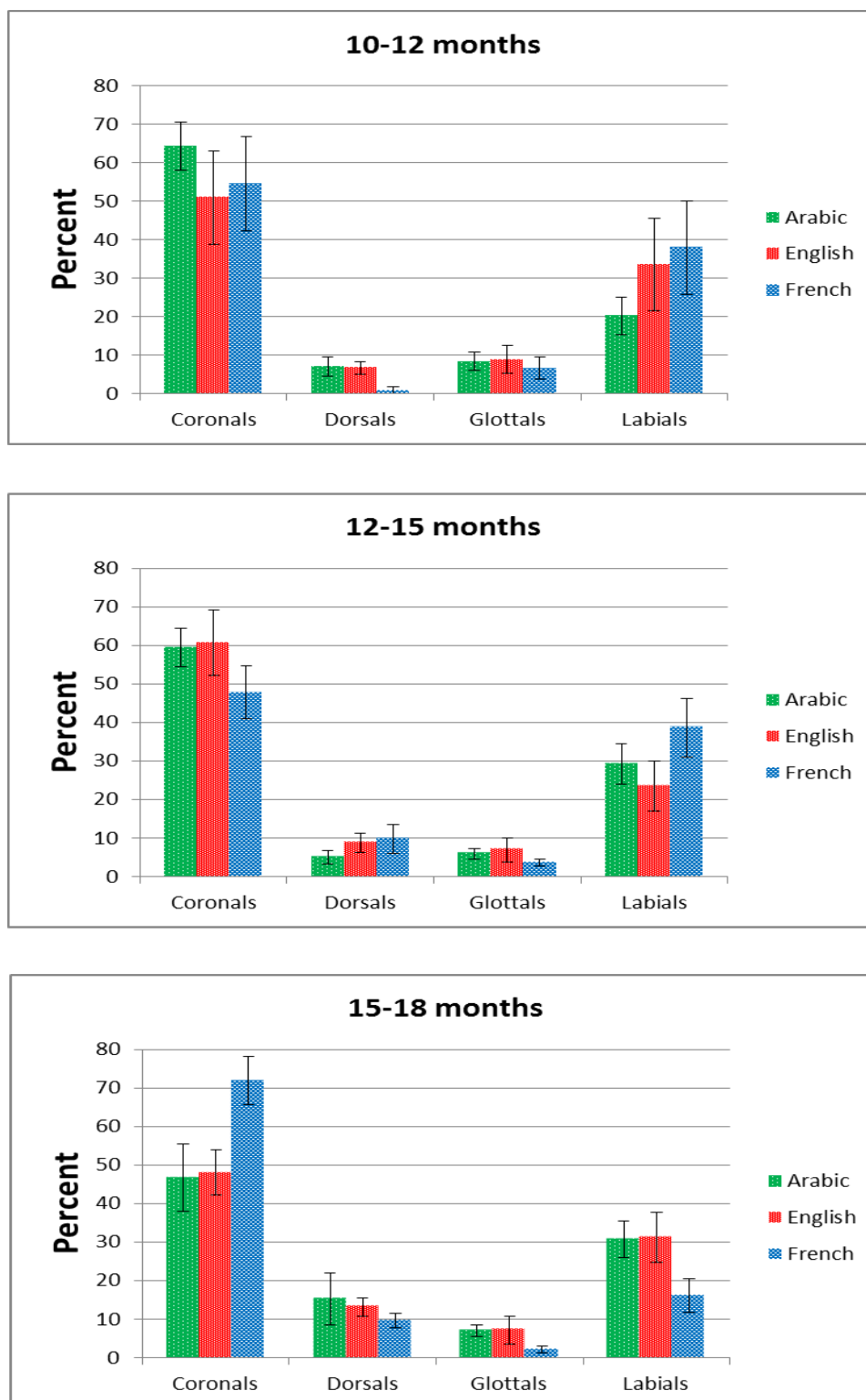


Figure 4.7. The proportion of each place category as plotted by language groups.

Table 4.7

*The results of ANOVAs examining the effects of age, language, and language-age interaction on the percentages of each place category of Arabic, English, and French infants from 10-18 months of age*

<b>Coronal</b>	
Age	$F(2,65) = 0.01, p = .99, \eta^2 = .00$
Language	$F(2,65) = 0.27, p = .77, \eta^2 = .01$
Age*Language	$F(2,65) = 2.41, p = .06, \eta^2 = .13$
<b>Dorsal</b>	
Age	$F(2,65) = 4.80, p = .011, \eta^2 = .13$
Language	$F(2,65) = 0.69, p = .50, \eta^2 = .02$
Age*Language	$F(2,65) = 1.16, p = .34, \eta^2 = .07$
<b>Glottal</b>	
Age	$F(2,65) = 1.01, p = .37, \eta^2 = .03$
Language	$F(2,65) = 1.81, p = .17, \eta^2 = .05$
Age*Language	$F(2,65) = 0.15, p = .96, \eta^2 = .01$
<b>Labial</b>	
Age	$F(2,65) = 0.36, p = .68, \eta^2 = .01$
Language	$F(2,65) = 0.30, p = .74, \eta^2 = .01$
Age*Language	$F(2,65) = 2.30, p = .068, \eta^2 = .124$

### 4.2.3 Results for manner of articulation

Two-way between-subject ANOVAs were conducted to examine the effects of age, language, and language-age interaction on the proportion of stops, nasals, fricatives, glides, and liquids of Arabic, English, and French infants from 10-18 months of age. Table 4.8 presents a summary of descriptive statistics for each manner of articulation category, and Figure 4.8 shows the proportion of each manner category as plotted by language groups across the three age groups (10-12, 12-15, and 15-18 months). The analyses show that no language, age, and language-age interaction effects occurred in the proportions of production in each manner category across language groups, except for a trend of age effects on the glides. The results of the ANOVAs are presented in Table 4.9. The Scheffe's test showed that no significant difference occurred in the means of the glides across the three age groups.

In summary, the results of this study indicated that using the phonetic transcription method revealed no significant crosslinguistic differences. In addition, age effects were not significant except for dorsals. Some phonetic categories showed non-significant changes (increase or decrease) with age across the language groups (Figures 4.7 and 4.8). For example, the proportions of coronals declined (except for French infants at 15-18 month), and the French infants showed a decrease in labials at 15-18 months compared to no changes among the Arabic and English infants. In addition, coronals and labials were preferred over the other place categories for all language groups. Similarly, stops and then nasals and glides were preferred over the other manner categories for all language groups.

Table 4.8

*The descriptive statistics in percentages for each manner category of Arabic, English, and French infants from 10-18 months of age*

<b>10-12 months</b>						
	Arabic		English		French	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Fricatives	13.69	9.26	11.75	10.88	9.14	7.34
Glides	18.30	12.44	25.50	25.62	25.16	11.55
Liquids	2.84	4.53	0.00	0.00	4.63	7.38
Nasal	20.18	24.03	21.57	22.30	20.73	22.45
Stops	44.99	24.48	41.18	31.69	40.34	10.12

<b>12-15 months</b>						
	Arabic		English		French	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Fricatives	11.48	6.39	17.14	13.57	16.41	17.05
Glides	15.56	10.80	16.17	12.01	16.19	8.11
Liquids	5.22	6.27	1.25	2.33	5.69	8.44
Nasal	10.10	8.79	8.87	6.66	26.13	11.45
Stops	57.64	12.24	56.58	21.54	35.59	20.81

<b>15-18 months</b>						
	Arabic		English		French	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Fricatives	10.78	5.30	9.08	10.16	10.07	9.71
Glides	13.82	7.95	20.23	12.36	10.88	7.09
Liquids	5.48	6.09	4.60	6.09	4.53	5.80
Nasal	19.48	13.16	13.17	10.00	17.11	15.75
Stops	50.44	16.57	52.92	19.64	57.42	17.57

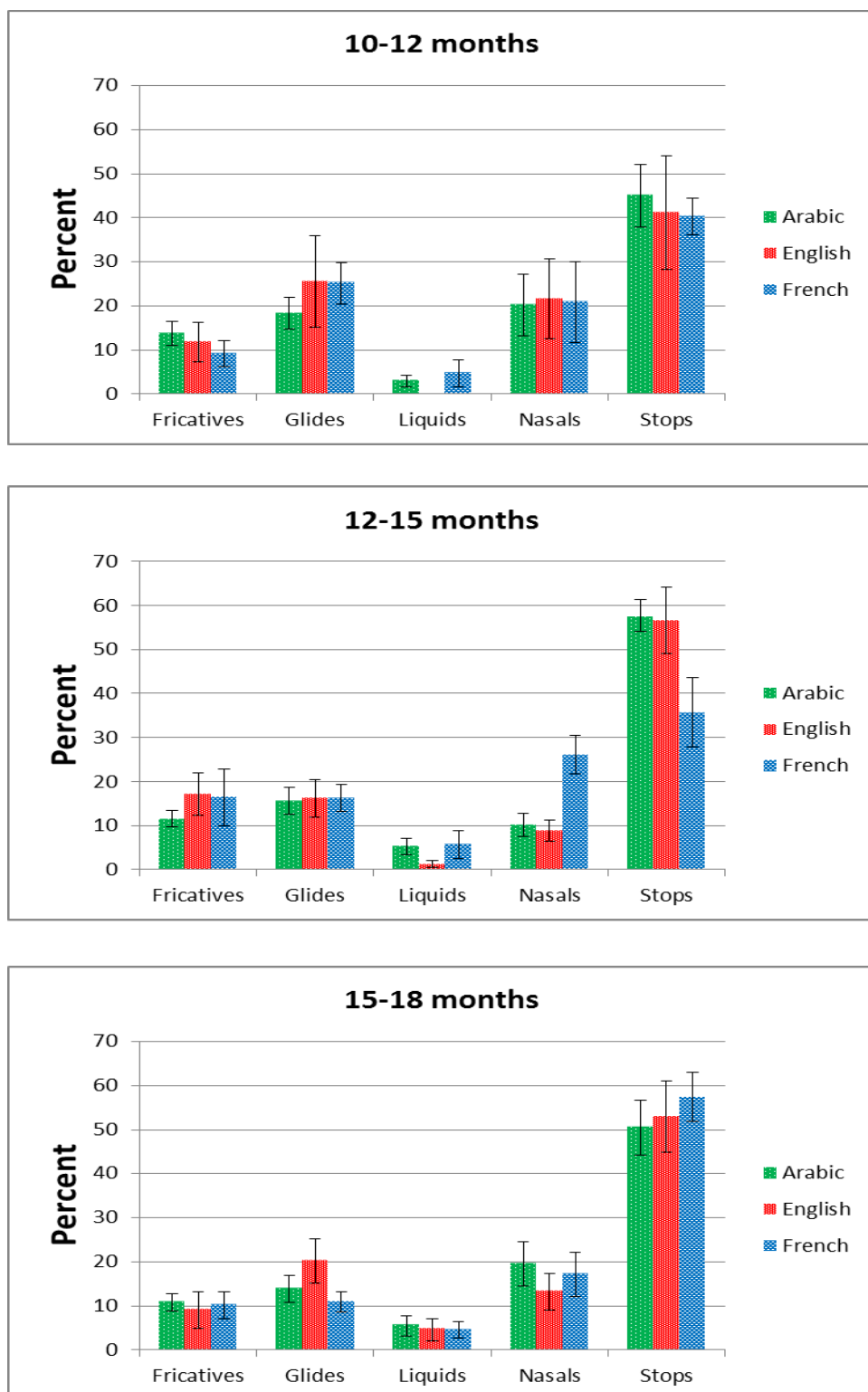


Figure 4.8. The proportion of each manner category as plotted by language groups.

Table 4.9

*The results of ANOVAs examining the effects of age, language, and language-age interaction on the percentages of each manner category of Arabic, English, and French infants from 10-18 months of age*

<b>Fricative</b>	
Age	$F(2,65) = 1.53, p = .23, \eta^2 = .05$
Language	$F(2,65) = 0.04, p = .97, \eta^2 = .00$
Age*Language	$F(2,65) = 0.64, p = .63, \eta^2 = 0.4$
<b>Glide</b>	
Age	$F(2,65) = 2.76, p = .07, \eta^2 = .08$
Language	$F(2,65) = 0.87, p = .43, \eta^2 = .03$
Age*Language	$F(2,65) = 0.59, p = .67, \eta^2 = .04$
<b>Liquid</b>	
Age	$F(2,65) = 0.98, p = .38, \eta^2 = .03$
Language	$F(2,65) = 1.69, p = .19, \eta^2 = .05$
Age*Language	$F(2,65) = 0.43, p = .79, \eta^2 = .03$
<b>Nasal</b>	
Age	$F(2,65) = 0.78, p = .46, \eta^2 = .02$
Language	$F(2,65) = 0.97, p = .38, \eta^2 = .03$
Age*Language	$F(2,65) = 0.96, p = .44, \eta^2 = .06$

## **5. DISCUSSION**

The purpose of this chapter is to present interpretations, explain the implications of the findings, and suggest directions for future research. Section 5.1 restates the research objectives. Section 5.2 summarizes and interprets the findings. Section 5.3 discusses the significance of the present work, including both theoretical implications and practical applications. Section 5.4 describes the limitations of the study. Section 5.5 includes some directions for future research. Section 5.6 presents a conclusion, which includes the main findings.

### **5.1. An overview of the research objectives**

As indicated earlier, the primary purpose of this study was to trace the speech development of Arabic infants from 10 to 18 months of age. Moreover, research describing the speech development of this population at this age range was lacking. The current investigation provided original data on Arabic babbling. The analyses of these data contribute to our knowledge about the course of Arabic speech development during infancy. Furthermore, the description of monolingual Arabic babbling in relation to the babbling of monolingual English infants and monolingual French infants enabled the examination of linguistic environmental influences on babbling. Such crosslinguistic investigations on the babbling of infants learning Arabic (Semitic language with a simple vowel inventory) and infants learning English or French (Indo-European languages with complex vowel inventories) contribute to the knowledge of the processes underlying infant speech development by showing unique patterns of changes in the babbling of Arabic infants compared to the other two language groups. The role of ambient language

input in infant early speech production—during the first 18 months of life—is a controversial research issue. One line of research claims that background language plays a minor role in infant babbling and that changes in babbling are due to the biological maturation of the vocal tract structures and oromotor control (e.g., Engstrand et al., 2003; MacNeilage & Davis, 2000, 2001; Oller & Eilers, 1982). Another line of research claims that ambient language input plays significant roles in the changes in infants' early speech productions, and that infants learning different languages babble differently (e.g., de Boysson-Bardies et al., 1989; Rvachew et al., 2008; Rvachew et al., 2006). The presence of crosslinguistic variations in babbling provides evidence for the early influences of ambient language input on the development of infant speech production and that the changes in infant babbling are not based merely on anatomical and motor control maturation, but rather on a more complex interaction of different factors including ambient language input.

This study hypothesized that infant age and language background influenced infant babbling. By using a cross-sectional quantitative causal-comparative research design (Gay et al., 2012), two studies were conducted. The primary study (Study One) involved the analyses of the acoustic vowel space of Arabic infants between the ages of 10-18 months, and then crosslinguistic comparisons of the acoustic vowel spaces of Arabic infants to those of English infants and French infants. The specific aims of this study were to describe the developmental changes in the acoustic characteristics of the vowel space of Arabic infants, and to identify crosslinguistic variations in infant vowels as measured by vowel space parameters (i.e.,  $M_{F1}$ ,  $M_{F2}$ ,  $SD_{F1}$ , and  $SD_{F2}$ ), vowel space features (i.e., extreme values of the acute, grave, diffuse, and compact features), and



vowel space area, which supports the *interactional hypothesis* (de Boysson-Bardies et al., 1989). Particularly, this study predicted that the infant vowel space is related to the complexity of the vowel system in the ambient language, and that infants learning a simple vowel system will achieve larger and earlier expansion of their vowel spaces as compared to infants learning a complex vowel system. The secondary study (Study Two) involved the analyses of the consonantal repertoires of Arabic infants between the ages of 10-18 months, and then a comparison of these consonantal repertoires to those of English infants and French infants. The specific aims of this study were to describe the developmental changes in the consonantal repertoires of Arabic infants, and to show that early crosslinguistic differences in the vowel space (Study One) were not accompanied by crosslinguistic differences in consonants as described by phonetic transcription.

## **5.2. A summary and interpretations of the findings**

As hypothesized in this study, the findings of the primary study (Study One) on the investigations of Arabic infants' vowel space showed supporting evidence for the *interactional hypothesis* (de Boysson-Bardies et al., 1989) and that Arabic infants' babbling showed age-related patterns of changes and language-specific patterns of changes in comparison to the babbling of English infants and French infants. In addition, the findings of the secondary study (Study Two) on the investigations of the consonantal repertoires showed no crosslinguistic differences in the manner and place categories, and so crosslinguistic variations in the vowel space (Study One) were not accompanied by crosslinguistic differences in consonantal inventories based on perceptual measurements (i.e., phonetic transcription).

This section begins with a brief recapitulation of the outcomes of Study One on the investigation of infants' vowel space and Study Two on the investigation of infants' consonantal repertoires. Then, a further discussion follows, which includes interpretations and explanations of the findings on Arabic babbling with respect to common and language-specific patterns of changes in the size and shape of the vowel space compared to the two language-comparison groups and the existing literature.

### **A summary of the findings**

Table 4.4 summarizes the findings of Study One on the crosslinguistic comparisons of infants' vowel space of the three language groups from 10-18 months of age. In addition, Figures 4.1 to 4.6 illustrate the presence of developmental changes common among the three language groups, and the language-specific variations in babbling. With respect to the vowel space parameters, the French F1 values showed a decline across the age range compared to the unchanged Arabic F1 and English F1 values, whereas the English F2 values decreased compared to the unchanged Arabic F2 and French F2 values. The analysis of the grave-acute feature showed that French *acuteness* values declined with age compared to the unchanged Arabic and English *acuteness* values, whereas a common decline in the *graveness* values was observed across the three language groups. In addition, the analysis of the compact-diffuse feature showed that the Arabic *compactness* values declined with age compared to the stable French and English *compactness* values, whereas the *diffuseness* value only increased among French infants. Furthermore, compared to the French and English infants, Arabic infants showed earlier expansion of vowel spaces at a younger age as indexed by a larger

vowel space area and extreme values of *diffuseness*, *graveness*, and *compactness* features—hence, a language-specific effect on the pace and the extent of vowel space expansion. In contrast, the findings of Study Two showed a similarity in the manner and place categories among the three language groups from 10-18 months of age.

### **Interpretations of the findings**

The analyses of the acoustic vowel space of the Arabic infants showed relatively unchanged F1 and F2 values from 10 to 18 months of age. Despite the stability in the center of the vowel space, Arabic infants appeared to expand the range of their F2 values with age—as indicated by a significantly increased  $SD_{F2}$  value, suggesting that they were expanding the range of horizontal tongue positions. The changes in the range of tongue advancement with age may account for the significant decrease in the *compactness* value, a significant decrease in the *graveness* value, and a non-significant increase in the *diffuseness* value (recall that the compact-diffuse dimension is related to the relative difference between F2 and F1, and the grave-acute dimension is related to the average of F1 and F2).

Considering the age variable, the patterns of changes in the vowel space of Arabic infants are suggestive of stability in the center of the vowel space due to unchanged F1 and F2 values, a great expansion toward the grave and compact corners of the vowel space, and less expansion toward the diffuse corner. The Arabic vowel inventory is characterized by a simpler triangular vowel space of three vowels (i.e., /i/, /u/, and /a/), their long counterparts (i.e., /i:/, /u:/, and /a:/), and two diphthongs (i.e., /aw/ and /ay/). It appears that Arabic infants keep the center of their vowel space unbiased toward any

directions during early infancy, since the three corner vowels may exert equal attractive power. Further, the vowel space expansion and the achievement of extreme values in diffuse, grave, and compact corners at a young age—compared to English and French infants—suggest that Arabic infants stretch their vowel spaces peripherally toward the corners of the ambient language input (i.e., the diffuse corner [i/], grave corner [u/], and compact corner [a/]) without any expansion toward the acute corner, which does not exist in the Arabic adult vowel space. In addition, the presence of /a/ in the two Arabic diphthongs may further account for the Arabic expansion toward the compact corner with no changes in the acute corner.

Furthermore, the Arabic infants exhibit age-related changes during the course of vowel space development. First, Arabic infants had a relatively larger vowel space at all ages compared to the French- and English-learning infants although some expansion in vowel space area was observed between 10 and 18 months (see Figures 4.5-4.6). Second, comparing the patterns of changes in the vowel space of Arabic infants to those changes in English and French infants illustrates the presence of developmental changes that are common among the three language groups. The developmental changes or age-related pattern of changes were observed in the expansion of the vowel space toward the grave corner. All three language groups showed a significant decrease in the *graveness* value with age, suggesting that they were producing more vowels in the high-back corner (i.e., /u/ corner in the adult vowel space). In addition, such common expansion may suggest that this vowel space region is first to reflect age-related changes toward achieving a vowel space with lower F1 and F2 values, which are characteristics of a mature vowel

space. Moreover, the three language groups showed an increase in the *diffuseness* value with age, which was only statistically significant for French, suggesting that the three language groups were producing more vowels in the high-front corner of the vowel space (i.e., /i/ corner in the adult vowel space). These age-related patterns are consistent with previous findings of restricted vowel space at a young age with predominate central and neutral vowels, and increasing production of high-front and back-front vowels as the infant matures (e.g., Buhr, 1980; Davis & MacNeilage, 1995; Kent & Bauer, 1985; Kent & Murray, 1982; Lee et al., 2010).

This study suggests that the vowel space stretches—in this case, a crosslinguistic expansion toward grave and diffuse corners—due to maturational changes of the vocal tract and oromotor control. As discussed in the introduction, infants start off with a small vocal tract and immature oromotor control; then, during their first 18 months of life, they experience a period of accelerated growth of their vocal tract combined with a gradual oromotor control over the articulators (Green et al., 2000; Green et al., 2002; Kent, 1992; Lieberman et al., 2001; Rommel et al., 2003; Siebert, 1985; Steeve et al., 2008; Thelen, 1991; Vorperian et al., 1999; Vorperian et al., 2005; Vorperian et al., 2009). By using an articulatory-to-acoustic model (the Variable Linear Articulatory Model [VLAM]), Menard et al. (2004) conducted an experiment to simulate the developmental changes in vowel space from infancy to adulthood. This model is based on previously published data on the growth of the vocal tract. The vocal tract length was set to five growth stages: 7.70 cm for a newborn, 10.67 cm for a 4-year-old, 12.65 cm for a 10-year-old, 15.36 cm for a 16-year-old and 17.45 cm for a 21-year-old. In addition, the model has seven articulatory parameters that simulate vocal tract structures (i.e., labial protrusion, labial aperture,

tongue tip position, tongue body position, tongue dorsum position, jaw height, and larynx height). These parameters were modified to produce different vocal gestures for vowel productions. The VLAM was set to simulate the maximum vowel space (MVS) that encompasses all possible vowels. The VLAM was able to produce the MVS at each of the five growth stages, which was confirmed by the perceptual study by native adult listeners, suggesting that infants with a small vocal tract were able to produce a great variety of vowels when they gain adult-like oromotor control (Menard et al., 2004). When the articulatory parameters of VLAM were set to simulate a vocal tract with a limited range of movements resembling infants' oromotor control, the simulated vowel spaces were small and limited with less expansion to the corners (Rvachew et al., 2006), suggesting a greater role for immature oromotor control on the production of different vowels during early infancy. In addition, Menard et al.'s (Menard et al., 2004; Rvachew et al., 2006) findings suggest that the short length of the vocal tract does not appear to prevent an infant from producing a great variety of vowels.

Second, a description of the changes in the vowel space of Arabic infants in relation to the vowel spaces of English infants and French infants reveals language-specific patterns of changes among the three language groups. The unique Arabic patterns of changes are found in the unchanged F1 and F2 values and a significant expansion toward the compact corner. More importantly, Arabic infants appeared to achieve a larger vowel space compared to the English and French infants as indicated by a larger size of the triangular vowel space and more extreme values in *diffuseness*, *graveness*, and *compactness* features from an earlier age of 10 months or perhaps younger than 10 months ( see Figures 4.2, 4.4-4.6 ). The speculated account of these

changes—as discussed previously—is due to the influences of the adult’s triangular vowel space of /i/, /u/, and /a/ that imposes equal forces, which causes stability in the center of the vowel space with a flexibility to expand toward the corners of the vowel space of the ambient language. The Arabic vowel system is simpler (fewer categories) and presents less distractions to an infant learning it, and subsequently, Arabic infants appear to have a clearer perceptual anchor in their system that enables the development of earlier articulatory representations of the three Arabic corner vowels (diffuse, grave, and compact). Again, the expansion to the compact corners is enhanced by the additional usage of /a/ in the two Arabic diphthongs (/aw/ and /ay/).

Similar to the Arabic infants, both the English and French infants showed language-specific changes in their vowel space. In contrast to the stability of the mean Arabic vowel with age, English infants showed significantly decreased F2 values and unchanged F1 values, whereas French infants showed significantly decreased F1 and unchanged F2 values. Unlike the Arabic simpler triangular vowel space with small vocalic sets, English and French have a quadrilaterally rich vowel space (12 English vowels with 5 diphthongs, and 15 French vowels). The presence of rich vowel systems may privilege English and French infants to experience a great variety of vocalic segments in speech input. Such rich vocalic experiences are speculated to aid the shift of the mean vowel space toward the native mature vowel space by showing an earlier reduction in the F1 or F2 values—compared to Arabic infants experiencing small vowel sets and stable F1 and F2 values. However, the complexity of the English and French vowel systems may account for the smaller vowel space of English and French infants as indicated by smaller size of triangular vowel spaces and less extreme values of corner

vowels, compared to Arabic infants learning a simple vowel system and showing a larger size of the vowel space and greater extensions into diffuse, compact, and grave corners at an earlier age

Furthermore, the differences in the pattern of shift to the adult's vowel space of the English and French infants (i.e., decreased English F2 with unchanged F1 values, and decreased French F1 with unchanged F2 values) may be due to the language-specific characteristics of the adult's vowel systems of the two languages. The differences in the infants' vowel space shift in relation to an adult's vowel space are not explicitly explained by the current study; however, speculative explanations can be drawn. For example, French has more oral vowels in the adult's /i/-/u/ dimension of the vowel space, a region characterized by vowels of lower F1 values. In addition, French oral vowels with lower F1 values occur more frequently than oral vowels with higher F1 values. Based on Malécot's (1974) report on the distributions of phonemes in the Spoken Paris dialect of French, the total occurrence of vowels in the examined corpus ranged from 16,216 to 787. Oral vowels with occurrences higher than 10,000 are /a/ (16,216), /e/ (16,051), /ə/ (10,891), and /i/ (10,797), suggesting that relatively more oral vowels with lower F1 values occur more frequently in the French spoken language (i.e., /i/, /e/, and /ə/ [mid vowel with lower F1 compared to vowels with high F1 such as /œ/]). Together, these two accounts (the relative frequent oral vowels in the French high-front area [/i/, and /e/] and the high proportion of oral vowels with lower F1 values in the French vowel space) may play a facilitative role in the French decreased F1 value. In addition, the two accounts may explain the unique French pattern of vowel space expansion to regions characterized



by low F1 values (i.e., diffuse and grave corners). Similar to French, the shift in the English vowel space with a decreased F2 value may be explained with respect to the frequency of the occurrence of phonemes. According to Mines, Hanson, and Shoup (1978), the most frequent vowels in spoken English—based on a total occurrence of English phonemes—is /ə/ including its rhotic variant (% 9.06), then followed by /i/ (% 3.69), /ɪ/ (% 3.64), and /ε/ (% 3.21). The greatest proportion of /ə/, which is characterized by a mid F2 value, may impose a decline on the F2 value of the mean English vowel.

To recap, the crosslinguistic variations in the vowel space are as follows. First, the unique Arabic patterns of changes are found in the unchanged F1 and F2 values, a significant expansion toward the compact corner, a larger size of the vowel space and more extreme values in *diffuseness*, *graveness*, and *compactness* features at 10 months of age. Second, the unique English pattern of changes was significantly decreased F2 values and unchanged F1 values with age. Third, the unique French pattern of changes are significant with respect to decreased F1 and unchanged F2 values, expansion toward the diffuse values, and shrinking from the acute corner toward the center of the vowel space.

The current description in the present study of Arabic babbling in relation to English and French babbling yielded outcomes that showed the presence of crosslinguistic variations in babbling—particularly in the infant vowel space—that contributes to the literature on ambient language influences on babbling. Such findings are in line with similar earlier acoustic investigations and provided evidence for the *interactional hypothesis* (de Boysson-Bardies et al., 1989). For example, de Boysson-Bardies et al.'s (1989) study showed crosslinguistic differences in the location of the

center of the infant vowel space of four language groups at 10 months of age. Compared to de Boysson-Bardies et al.'s study, the current study showed crosslinguistic differences in the pattern of changes in the center of the vowel space. As age increased, the center of the Arabic infants' vowel space was stable compared to the English infants' vowel space that showed a shift with a decreased F2 value, and to the French infants' vowel space that showed a shift with a decreased F1 value. In addition, the current study traced ambient language influences across a larger age range (10 to 18 months) with a larger sample size as well. Furthermore, Rvachew et al. (2006) conducted a crosslinguistic comparison of the vowel space of English and French infants. These two infant groups formed the comparison groups for the Arabic group in the current study. The French infants and English infants were not compared directly with each other in the current study; rather, the Arabic infants were compared to the French infants and then to the English infants separately. The current study showed that the vowel space of Arabic infants learning a simple vowel inventory is stretched at a younger age—hence, a larger size—compared to English and French infants learning complex vowel inventories. Moreover, Rvachew et al. (2008) showed crosslinguistic differences in the extreme grave vowels of English infants and French infants from 8 to 18 months. The frequency of grave vowels judged as /u/ by native listeners was greater among English infants than among French infants. The current study showed a unique Arabic expansion into the /a/ region as indicated by a decrease in the Arabic *compactness* value compared to no changes among the French and English infants. In addition, the current study showed that Arabic infants achieved a larger size of triangular area of the vowel space and more extreme values in *diffuseness*, *graveness*, and *compactness* features at 10 months of age. Overall, the findings of the

current study are consistent with other investigations that report crosslinguistic differences in different aspects of babbling (e.g., prosodic features), and thus produce positive evidence for the *interactional hypothesis* (e.g., de Boysson-Bardies et al., 1984; de Boysson-Bardies et al., 1986; de Boysson-Bardies & Vihman, 1991; Hallé et al., 1991; Lee et al., 2010; Levitt & Utman, 1992; Levitt & Wang, 1991). However, by using an infraphonological description, a phonetic transcription, and a detailed acoustic description of the infant vowel space, the current study also presented additional evidence about a language-specific pattern of changes from a large scale sample of Arabic infants learning a simple vowel inventory and through the crosslinguistic comparisons of two infant language groups learning either English or French (languages with complex vowel inventories). The current study provided evidence that the expansion in infant vowel space is related to the complexity of the vowel system of ambient language (i.e., learning a simple vowel system is expected to make the task to expand the vowel system easier).

The methodology that previous research employed to conclude that ambient language did not influence infant babbling may have played a role in arriving at such a conclusion. Studies that showed these kinds of negative findings may be categorized as descriptive reviews of previously published studies that used transcription analyses of precanonical sounds and canonical sounds (e.g., Locke, 1983), transcription studies based on infraphonological description (e.g., MacNeilage & Davis, 2000, 2001; Oller & Eilers, 1982), and perceptual studies (e.g., Atkinson et al., 1968; Engstrand et al., 2003; Thevenin et al., 1985). A comprehensive evaluation of these studies is found in the literature review section of the current study. An overall observation about these scholarly investigations suggests that the use of auditory-based measurements such as

transcription and listening tasks—among other variables (e.g., a lack of statistical analyses and a small sample size)—may have yielded negative conclusions about the early effects of ambient language input on babbling. The difficulty in recovering early language-specific changes in consonants using the transcription tool is highlighted by the results of Study Two in the current investigation. The outcomes of Study Two are discussed next.

The phonetic transcription analyses conducted in Study Two described the consonantal repertoires of Arabic infants between the ages of 10-18 months—based on the manner and place of production. With respect to place of articulation, the coronal was the largest category across the three age groups (64% at 10-12 months, 60% at 12-15 months, and 47 % at 15-18 months), which was predominated by /d/, /t/, /n/, and /j/ with fewer other sounds such as /ʒ/, /l/, /ð/, /θ/, and so on. The second largest category was labial across all three age groups (20% at 10-12 months, 30% at 12-15 months, and 31 % at 15-18 months), which was predominated by /b/, /p/, /w/, and /m/. The glottal and dorsal categories were produced less frequently across the age ranges. The glottals were predominated by /h/ and a few instances of /ʕ/ and /ħ/; whereas, the dorsals were predominated by /k/ and /g/ with fewer instances of /ŋ /, /x/, and /ɣ/. With respect to the manner of articulation, the stops were the largest category across the three age groups (45% at 10-12 months, 58% at 12-15 months, and 50 % at 15-18 months), which was predominated by /p/, /b/, /d/, /t/, /k/, and /g/. Nasals, glides, and fricatives followed with proportions ranging from 10% to 20% (the proportion of fricatives represented 11% to 14%). The proportion of liquids, which were predominated by /l/, showed a minimum

increase of 3% at 10-12 months to 5.5 % at 15-18 months. The findings from consonantal descriptions of Arabic babbling with respect to the predominant phonetic categories were consistent with the previous findings discussed in the literature review section on the consonantal repertoire of infants learning languages other than Arabic (e.g., Kent & Bauer, 1985; Locke, 1983).

Unlike the acoustic comparisons of the vowel space, the findings from Study Two on the crosslinguistic comparisons of consonantal repertoires of Arabic infants to English infants and French infants showed a similarity in the manner and place categories among the three infant groups from 10 to 18 months. For example, the coronals and labials were preferred over the other place categories among the three language groups; similarly, stops and then nasals and glides were preferred over the other manner categories. A few speculations can be drawn about the lack of crosslinguistic variations in the proportions of phonetic categories. For example, the language-specific changes in babbling are subtle and may not be observed when using only auditory-based measurements (e.g., IPA). In the current study, the weak sensitivity of the auditory-based measurements to recover the early effects of ambient language on babbling was not problematic nor an objective, since different methods were used to analyze different aspects of babbling (i.e., an acoustic analysis of vowels and a phonetic transcription for consonants). Hence, no evidence was provided for the sensitivity of auditory-based measurements to track any early ambient language changes. In addition, the accurate production of many consonants is acquired over several years compared to early vowel acquisitions (Sander, 1972; Stoel-Gammon & Pollock, 2008). Furthermore, a high individual variability exists in consonantal acquisition in infants exposed to the same language (Ferguson & Farwell, 1975; Vihman

& Croft, 2007). Infants tend to produce sounds differently across words they attempt. For example, Ferguson and Farwell (1975) observed some variability in the phonological development of three English children from age one year until the production of their first 50 words. Their findings showed a great variability among the three subjects. For instance, subject T showed a preference for using sibilant fricatives, affricates, and a velar stop in her first 50 words, whereas the other two children showed a minimal use of these phonemes, suggesting late acquisition compared to subject T. In addition, such variability also has been observed at the within-child level. For example, subject K showed about 10 different productions of the word “pen” in a single 30-minute recording session (e.g., [hɪn], [bɑ<sup>h</sup>], [p<sup>h</sup>ɪn], [d<sup>h</sup>au<sup>N</sup>]). Thus, the ambient language influences were not evident in the consonantal inventories as described by phonetic transcription.

The presence of crosslinguistic variations in babbling suggests that vocal development in infancy is a dynamic process. Such a process includes internal sources—for example, anatomical and speech motor variables—that impact developmental changes in babbling, and external sources—for example, environmental input to the infants such as linguistic background, visual information, and social interactions—that impact language- or environment-specific changes.

The environmental influences in the form of linguistic input are not only due to differences in the phonemic inventory (types of vowels and consonants in the ambient language), but also are due to the crosslinguistic differences in the input frequency (i.e., the frequency of the occurrence of phonemes in the input or the segmental distributions) and the functional load of the segments (i.e., the importance of phonemic contrasts within

the phonological system of a language for distinguishing words). The low frequency segments in a language are acquired late, whereas the high frequency segments are acquired early (Macken, 1995), and phonemes with high functional loads are acquired first—the higher the occurrence of a phoneme in minimal pairs, the higher the functional load is (Ingram, 1989; Stokes & Surendran, 2005). By analyzing previously published acquisition data, Stokes and Surendran (2005) examined the relationship among articulatory complexity, input frequency, and the functional load of consonants with respect to English (8-26 months), Cantonese (15-30 months), and Dutch (16-48 months) infants. The input frequency and the functional load of consonants were obtained from the published corpus for the three languages. The functional load for each consonant was defined as “the weighted sum of the functional load of the binary oppositions between it and other phonemes with which it is likely to merge [due to sharing place of articulation, manner of articulation, or/and voicing]” (Stokes & Surendran, 2005, p. 581). The value for articulatory complexity was set for each sample according to Kent’s motoric analysis of the consonant acquisition (discussed in the introduction of the current study). The findings showed that articulatory complexity was correlated with the age at which consonants emerged in the Cantonese and English data, and with the accuracy of consonants production for the English and Dutch data, suggesting that children tend to produce first consonants with the least articulatory complexity. Further, input frequency was strongly correlated with the age at which consonants emerged in the Cantonese data and with the accuracy of consonants production for the Dutch data, suggesting that the Cantonese and Dutch infants tended to first produce high frequency consonants. In addition, functional load was moderately correlated with the consonant emergence in the

English data, suggesting that the English infants tended to first produce consonants with a higher functional load. Similarly, Pye, Ingram, and List (1987) showed the significance of input frequency on the rate of the consonant acquisition of English children (1;5 to 2;2 years) and Quiche children (1;7 to 3;0 years). The frequency of occurrence of English consonants was estimated from the published data for older children (5; 10 to 8; 4 years), and for the Quiche children was estimated from the adult forms that they attempted to produce. The findings showed a strong correlation of the order of acquisition to the frequency of occurrence ( $r = .76$  for the Quiche data and  $r = .55$  for the English data). For example, affricate /tʃ/, which was the most frequent Quiche sound, was acquired earlier among Quiche children compared to the late acquisition of /tʃ/ in English, which is one of the least frequent English sounds. Although their findings are based on older infants and toddlers, the studies of Stokes and Surendran (2005) and Pye et al. (1987) showed the important contribution of the input frequency and the functional load of a language in early phonological development.

However, the input frequency of segments in adult speech or adult-directed speech (ADS) may not present similar properties of language input with respect to infant or infant-directed speech (IDS). As discussed in the introduction of the current study, based on an acoustic analysis of the F1 and the F2 values, Kuhl et al. (1997) reported greater maternal vowel space in the IDS compared to the ADS. In addition to simpler and shorter IDS utterances, Lee and her associates (Lee, Davis, & Macneilage, 2008; Lee et al., 2010) reported differences in the segmental distribution in the ADS and IDS. For example, Korean IDS used fewer coronals, glottals, and fricatives but more labials compared to Korean ADS; and Korean IDS used more mid-central and low-central



vowels, whereas Korean ADS used more mid-front and high-central vowels (Lee et al., 2008). In addition, English IDS showed greater proportions of high-back vowels compared to English ADS (Lee et al., 2010). Accordingly, infants are not expected to match the ADS; rather, they are expected to match the inputs in the IDS. Lee et al. (2010) reported that the higher proportion of /u/ in the English sample compared to the French sample in Rvachew et al. (2008) is consistent with their findings of frequent /u/ in the English IDS but not in the English ADS; and that IDS is the frequent ambient language input for infants. Therefore, ambient language inputs provided to infants in the form of IDS are not the same as ADS. Accordingly, infants are not expected to match the ADS; rather, they are expected to reflect some IDS characteristics in their babbling while they cope with the constraints imposed by their maturing articulatory systems.

The other environmental inputs that contribute to changes in speech production during infancy include social input from social feedback for vocalizations. The active communicative relationship between mothers and their infants acts as a facilitative function or scaffolding for normal speech acquisition (Goldstein & Schwade, 2009; Kuhl, 2003). Infants tend to change their babbling as a result of maternal social feedback (e.g., smiling, touching). In a vocal learning facilitation situation, Goldstein and his associates (Goldstein, King, & West, 2003; Goldstein & Schwade, 2008) examined the role of social interaction between infants (8 to 9 months) and their mothers. Infants and their mothers were divided into two groups: 1) infants in the contingent condition received social feedback immediately after each babbling, such as smiling, moving closer, and touching with maternal vocal models of fully-resonant vowels or words, and 2) infants in the control condition received random social feedback with maternal vocal models.

Infants in the contingent condition increased their babbling and incorporated maternal vocal models in their babbling by producing more adult-like syllables compared to the infants in the control condition who showed no changes in their babbling (Goldstein & Schwade, 2008). In addition, the infants in the contingent condition showed changes in their babbling (e.g., increased the proportion of mature babbling) as a result of social feedback only, without vocal models (Goldstein et al., 2003). Thus, Goldstein et al.'s findings showed that the speech input presented in social interactions caused changes in infant vocalization, which suggests that infants are sensitive to input from their environment—in this case, to social input. The findings from the current study showed that infants exposed to different languages babble differently, suggesting infant sensitivity to environmental linguistic input, which is in line with Goldstein et al.'s findings on infant sensitivity to social input.

Furthermore, environmental influences in the form of visual and auditory input cause changes in infant vocalizations. Infants are able to extract phonetic information presented in integrated audio-visual speech input. In laboratory settings, infants were familiarized with the visual stimulus of a talker producing /a/ or /i/ without sound, and then in the testing phase, they were presented with two video clips of the same talker articulating both sounds synchronously, but only one of the two sounds (either /a/ or /i/) was played back. Infants (2-4.5 months) looked longer at the talker's face that matched the auditory stimulus, suggesting that the infants related the visual information with the auditory input for speech perception (Kuhl & Meltzoff, 1982; Patterson & Werker, 1999, 2003). Furthermore, audio-visual speech input appears to influence infant vocalizations. For example, many infants recruited in the study conducted by Patterson and Werker

(1999) imitated the vowels that they heard and matched to the talker's face. In addition, Kuhl and Meltzoff, (1996) played 5-minute segments of videotaped stimuli for 12-, 16-, and 20-week old infants. The stimuli included /i/, /a/, or /u/. Their findings showed that infants imitated the presented stimuli. For example, infants who heard /i/ produced significantly greater proportions of /i/-like vowels; and similarly, for the /a/ and /u/. Taken together, ambient environmental input in the form of audio-visual speech influences infant vocalizations; even a short-term exposure to linguistic input caused a change in babbling. Consistent with the just-cited studies, the current findings showed that language-specific exposure—occurring over the first few months of life—caused changes in infant babbling as indicated by the crosslinguistic variation in the infant vowel space.

The current study showed that infants' vocalizations are influenced by the speech input provided to them by the ambient linguistic environment—particularly crosslinguistic variations in the center and expansion of the vowel space. In addition, it showed that each of the three infant language groups presents a unique pattern of changes across the age range. Based on such findings, an assumption can be made that infant babbling *drifts* toward her/his household language background. However, the notion of *drifting* in babbling—as it first appeared in Brown's (1958) book—is not well defined in the literature. For example, what is the pattern of *drifting* in babbling? Do infants *drift* slowly and steadily toward ambient language or do they start with passive vocalizations unbiased by their ambient language input, and then at one point during the second 6 months of their life start producing acoustic-motor matches of their surrounding language? What is the starting point for such *drift* for infants exposed to different

languages? The current study does not provide answers to such questions; however, it raises these questions for future investigations and provides an alternative account of the role of ambient language input in babbling based on the *interactional hypothesis* (de Boysson-Bardies et al., 1989), and considering the active role of the infant language learner with her/his attentional preferences as shaped by the complexity of the input. Given that infants start off with an immature articulatory system (i.e., a small vocal tract and imprecise oromotor control), infants who are born to different linguistic communities start at the same semi-random place with respect to speech production and perception (e.g., language-general perception, similarity in precanonical babbling). At about 6 to 8 months of age, their perception is reshaped by the ambient speech input to become language-specific. During the first 8 to 10 months of life, infants have exclusively specific linguistic listening experiences. Such specific speech input imposes gradual influences on infant speech production, which increases in pace during the second year of life. Evidence for the slow language-specific changes comes from subtle crosslinguistic variations of babbling that need sensitive measures—such as acoustic analyses—to recover. These language-specific changes on infant speech production are caused by the interactions of infant speech perception, attentional biases, limitations of speech motor control, small vocal tract, and speech input (including ambient language and infant self-speech). The current study did not attempt to provide evidence for a *drift*; rather, it showed evidence for the influences of specific features of ambient language on babbling (e.g., how the complexity of the vowel system of ambient language impacts the pace and extent of expansion of the vowel space). Given that infants are exposed to the IDS form of their household language, the comparison of IDS and infant babbling should enable a

tracing of any direct connections between them and help to explain the *drifting* pattern. Still, the discovery of the pattern of the *drift* toward ambient language awaits future investigations that will focus on understanding the nature of infants' experience of ambient language input.

Overall, based on the current literature review and current findings, a brief summary can be put forward to describe the process of infant speech production and perception in relation to ambient language influences. The nature of speech input to the infant (e.g., IDS and language-specific input) and the infant's ability both to access and process that input (e.g., lack of hearing impairment) play an important role in the development of normal speech perception and production skills (e.g., canonical babbling and the production of meaningful speech). In addition, ambient language input exerts early influences on both infant perception and production. Infants rely on their statistical learning mechanism to decode the regularities in the distributions of speech inputs, which provides them with the capability to distinguish native and non-native speech sound contrasts. Infants experience developmental changes in speech perception toward language-specific perception; and ambient language input imposes a reorganization of infant perception, which causes them to perform well in the discrimination of native phonetic contrasts compared to non-native contrasts. Throughout the first year of life, infant vocalization is constrained by a small vocal tract and immature oromotor control of the articulators. Such limitations cause infant babbling to show a small and limited vowel space, and the production of few consonants and few word shapes. As the vocal tract increases in size and length and the infant gains better control over the articulators, the infant vowel space expands to its corners, and the phonetic inventory increases in size

with more syllable shapes. During this growth period, infants are able to modify their speech production and maintain sufficient control over the articulators. As proposed by Callan et al.'s (2000) research based on the DIVA model, the auditory, tactile, and proprioceptive feedback generated by the articulatory movements during early infant vocalization enables infants to develop an internal mapping between acoustic goals and vocal tract shapes. The acquisition of stable acoustic-phonetic matches enables infants to update such mapping to accommodate the maturational changes of their vocal tract morphology. In addition to developmental changes, ambient language input impose language-specific influences on infant speech production that have been observed before the end of the first year of life. For example, the current study presented additional evidence gleaned from the description of Arabic babbling in relation to English and French babbling, which showed crosslinguistic variations in the vowel space between the ages of 10-18 months.

### **5.3 Implications**

The findings about the influences of early ambient language on babbling have theoretical implications. Theories and models that attempt to describe and/or simulate infant vocalization need to consider that changes in infant babbling are due to a complex interaction of many variables, among them, maturational changes in the vocal tract and oromotor control, and linguistic environmental input. Such a claim is in accord with de Boysson-Bardies et al.'s (1989) *interactional hypothesis* that argues for the involvement of both biological and perceptual mechanisms in the production of infant speech, that suggests that infants learning different languages babble differently. Examples of the

biological-based models include the Frame/Content Theory (MacNeilage & Davis, 2000, 2001) that puts forward a biomechanical approach to explain infant babbling. As discussed in the introduction of the current study, infant vocalizations are the result of phonation with alternative opening-closing cycles of the mouth dominated by jaw involvement with a minimum contribution of the tongue and lip. The products of such rhythmic movements are CV-combinations with limited sets of vowels and consonants, which some scholars accept as contradiction to early role for ambient language on babbling. Based on current findings and other similar findings on crosslinguistic difference in babbling, Frame/Content Theory should consider that the CV vocalization reflects ambient language effects, which can be related to the prosodic characteristic, consonantal inventory, and vowel space. Another example of a biological model is Locke's proposed model of three stages—pragmatic, cognitive, and systemic—which correspond to the prelinguistic, the production of the first word, and the production of first 50-words periods, respectively. According to Locke (1983), only at the systemic stage does infant speech production move “in the direction of [the] adult system in ways it would not be expected without environmental stimulation” (p. 98). However, based on the findings of the current study and other similar studies, “environmental stimulation” also plays an early role in infant speech output during what Locke calls the pragmatic and cognitive stages.

The practical implications of the current findings include an emphasis on the significance of the quality of infants' experience with ambient language input for normal speech development. The variation in infant speech is accounted for by many variables, mostly biological and environmental. A common influence of speech input on babbling is

that it imposes a universal course for the normal development of prelinguistic stages across languages (i.e., the presence of universal milestones for vocal development). With respect to the vowel space, the universal effect is that infants—regardless of the ambient language—try to expand their vowel spaces. Furthermore, language-specific input cause changes in babbling that are part of the normal development process, but not universally. As shown in the current study, the learning of a simple vowel system yielded an early expansion of the vowel spaces of Arabic infants compared to English and French infants, which suggests that the more complex inputs to the French- and English-learning infants complicates their efforts to expand the vowel space toward the periphery. As such, the current study highlights the importance of encouraging parents to provide infants with rich-linguistic input even before they are capable of producing meaningful speech. Different biological and social factors (e.g., hearing impairment and the lack of child-mother interaction) may affect the quality of an infant’s experience with ambient language, and contribute to individual differences in speech and language development within and across linguistic communities. Furthermore, the role of enhanced input during early development to prevent or treat speech and language disorders is worthy of further research. A review of perceptual training research on children with speech sound disorders can be found in Rvachew and Brosseau-Lapr  s’s (2010) chapter.

## **5.4 Limitations**

Some limitations as to the internal validity of the current study exist. The unbalanced sample size among comparison groups (Arabic N= 31, French N= 23, and English N=20) is not ideal but occurred because the English and French participants



came from an existing dataset; all infants who met the inclusion criteria were included to have the participants evenly spaced in the 10- to 18-month-age range. In addition, the recording session for each Arabic infant was collected over the summer period because of time constraints and the parents' availability. Thus, this single recording session only provided access to infant babbling at one point in time and so, it may not be representative of the infant's typical vocalizations. Moreover, the majority of Arabic infants were screened for middle ear function but not for hearing sensitivity because of a lack of access to hearing assessment tools (the Arabic data collection was done abroad). The information about the infants' hearing sensitivity was based on parental reports. In addition, a comparison of the Arabic findings with other published studies of Arabic infants was not possible because developmental investigations that use acoustic analyses of vowels produced by Arabic infants do not exist. Furthermore, a good quality recording of the mothers' voices of the infants recruited in this study were not obtained. Such recordings would allow for an acoustic analysis of IDS and whether infant vowels are linked to IDS vowels. Although the speech samples were collected following the same protocol, some unavoidable issues occurred. Most recordings were done with one parent; however, that parent could be the mother or father, and a few recordings were done with a familiar person (a close relative). Even though the parents were instructed to speak as they usually spoke to the infant so to produce a spontaneous speech sample, the recording environment (e.g., the infant wearing a vest with a microphone clipped on, an observer in the same room, and recording equipment) may have affected the usual parent-infant interaction. Moreover, the multiple regression analyses of cross-sectional data do not permit identification of a specific age point where the language effect begins; however,

such information can be estimated from an examination of the regression lines. Last, the outcomes of the ANOVA on the crosslinguistic differences on the size of the vowel space should be considered with caution. The area sizes were strongly correlated with the number of vowels in the French sample ( $r = .78, p < 0.001$ ), but they were not correlated for the Arabic sample ( $r = .27, p = 0.15$ ) and the English sample ( $r = .3, p = 0.2$ ). In addition, the Levene's test for the homogeneity of variances of DV across groups for one-way ANOVA at 12-15 age interval was significant ( $F = 4.614, p = 0.02$ ), suggesting that the variances were not equal. Therefore, the significant effects resulting from the ANOVA test were considered at a lower alpha level (i.e.,  $p < 0.01$ ). Confidence in this finding would be enhanced by replication in a study in which the size of the speech samples per infant was controlled (See Molemans, Van Den Berg, Van Severen, & Gillis, 2012, for a discussion of speech sample size issues).

## 5.5 Future directions

Given that infants are typically exposed to their ambient language in IDS form, it would be worth doing an acoustic investigation of the maternal vowel space and compare it to infant vowel space. Such an analysis may provide a better explanation of the crosslinguistic variation in babbling. In addition, the current investigation can be extended to other important dependent variables; for example, do crosslinguistic differences exist in the prosodic characteristics of the babble produced by monolingual infants exposed to a different language background, as revealed by (a) measures of vowel duration in utterance internal and utterance final positions and (b) measures of utterance length in syllables? Furthermore, the possibility also exists for other researchers to extend

the current study to a different language group; such an investigation could potentially reveal language-specific patterns of changes in the vowel space that are comparable to the current findings for the three infant languages. Moreover, this study utilized a cross-sectional research design, and so a longitudinal investigation could potentially provide different informative data on ambient language effects and on infant maturational changes with age.

Furthermore, the observed crosslinguistic differences in infant vowel productions— particularly, the larger vowel space of Arabic infants due to the simpler Arabic vowel space compared to the other language groups—calls for crosslinguistic research on speech perception involving Arabic infants. Such research would enable the exploration of the link between the observed production differences and perception. The natural referent vowel (NRV) hypothesis posits that infants are born with perceptual biases for peripheral vowels versus less peripheral vowels in the F1/F2 space, i.e., directional asymmetries in vowel discrimination (Polka & Bohn, 2011). For example, 6- to 12-month-old Danish infants were significantly more successful in signaling changes in native and non-native contrasts from less peripheral vowels to more peripheral vowels (e.g., from /ɛ/ to /e/ or from /ʌ/ to /a/) compared to changes in more peripheral vowels to less peripheral vowels (e.g., from /e/ to /ɛ/ or from /a/ to /ʌ/). Further analyses showed that the asymmetry in vowel discrimination was maintained only for non-native contrasts for 10- to 12-month-old Danish infants and Danish adults compared to younger infants (6- to 9-month-old) who showed asymmetry patterns of perception for native and non-native contrasts alike. These findings are suggestive of language-specific influences on the universal perceptual bias (see Polka & Bohn, 2011, for a review of directional

asymmetries in vowel discrimination). Based on the NRV predications, future research is warranted to explore whether the bias for peripheral vowels is more robust in Arabic infants due to their simpler vowel space; a robust perceptual bias for these corner vowels might explain the rapid and early expansion of the productive vowel space in Arabic vocal development. Rvachew et al. (2008) have previously discussed the link between listening preferences for /u/ versus /y/ in English infants and the greater frequency of /u/ vowels in English versus French babble.

Finally, it is clear that a great need exists for additional research on the acquisition of Arabic phonology, particularly during the first years of life, since the data collection and analysis of infant data are more challenging compared to older populations. The relatively large dataset of analyzed Arabic infant babbling generated by this study provides a steppingstone to conduct further studies on Arabic phonological development.

## **5.6 Conclusion**

The present study focuses on describing the babbling of monolingual infants exposed to the Arabic language in Saudi Arabia. The description of Arabic babbling enables an investigation of the influences of ambient language input on the babbling of monolingual Arabic, French, and English infants from 10 to 18 months old. The findings show developmental changes in one corner (high-back/grave) of the vowel space, and language-specific changes in the center of the vowel space, the other three corners of the vowel space, and the triangular vowel space area. More importantly, the findings confirm the hypothesis that the developmental expansion of the infant vowel space is related to the complexity of the vowel system of the ambient language. As shown in the current

study, Arabic infants learning a simple vowel system had a larger size of the vowel space compared to French and English infants learning more complex vowel systems. In addition, the crosslinguistic difference in the vowel space was not accompanied by differences in the consonantal repertoires of the three language groups based on phonetic transcription analyses. The overall findings of the current study provide evidence for an *interactional hypothesis* and that the processes underlying infant babbling are influenced by a complex interaction of the biological development of the vocal tract and the input of ambient language.

## REFERENCES

- Abdala, C., & Folsom, R. C. (1995). The development of frequency resolution in humans as revealed by the auditory brain-stem response recorded with notched-noise masking. *Journal of the Acoustical Society of America*, 98(2 Pt 1), 921-930.
- Abdo, D. A. (1969). *Stress and Arabic phonology*. Thesis, University of Illinois, Urbana, Ill.
- Abu-Mansour, M. (2007). Meccan Arabic. In K. Versteegh (Ed.), *Encyclopedia of Arabic language and linguistics* (Vol. 3, pp. 179-189). Leiden; Boston: Brill.
- Al-Ani, S. H. (1970). *Arabic phonology: An acoustical and physiological investigation*. The Hague: Mouton.
- American Speech-Language-Hearing Association. (1997). *Guidelines for audiologic screening [Guidelines]*. Retrieved from [www.asha.org/policy](http://www.asha.org/policy).
- Atkinson, K., MacWhinney, B., & Stoel, C. (1968). *An experiment on the recognition of babbling*. Paper presented at the Language-behavior research laboratory working paper No. 15, Berkeley: University of California.
- Avaaz Innovations Inc. (1999). Time frequency response (Version 2.1). London, ON.
- Bibas, A. G., Xenellis, J., Michaels, L., Anagnostopoulou, S., Ferekidis, E., & Wright, A. (2008). Temporal bone study of development of the organ of Corti: Correlation between auditory function and anatomical structure. *The Journal of Laryngology and Otology*, 122(04), 336-342. doi: doi:10.1017/S0022215107006548
- Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer (Version 5.2.27). Retrieved from <http://www.praat.org/>
- Bredberg, G. (1967). The human cochlea during development and ageing. *The Journal of Laryngology and Otology*, 81(7), 739-758.

- Brown, R. (1958). *Words and things*. Glencoe, Ill.: The Free Press.
- Buhr, R. D. (1980). The emergence of vowels in an infant. *Journal of Speech & Hearing Research*, 23(1), 73-94.
- Callan, D. E., Kent, R. D., Guenther, F. H., & Vorperian, H. K. (2000). An auditory-feedback-based neural network model of speech production that is robust to developmental changes in the size and shape of the articulatory system. *Journal of Speech, Language, and Hearing Research*, 43(3), 721-736.
- Carter, M. G. (2001). Arabic. In J. Garry & C. R. G. Rubino (Eds.), *Facts about the world's languages: An encyclopedia of the world's major languages, past and present* (pp. 23-27). New York: H.W. Wilson.
- Chahal, D. (2007). Intonation. In K. Versteegh (Ed.), *Encyclopedia of Arabic language and linguistics* (Vol. 2, pp. 395-401). Leiden; Boston: Brill.
- Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Naatanen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*, 1(5), 351-353.
- Davis, B. L., & MacNeilage, P. F. (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research*, 38(6), 1199-1211.
- de Boysson-Bardies, B. (2001). *How language comes to children: From birth to two years*. Cambridge, Mass.; London: MIT Press.
- de Boysson-Bardies, B., Halle, P., Sagart, L., & Durand, C. (1989). A crosslinguistic investigation of vowel formants in babbling. *Journal of Child Language*, 16(1), 1-17.
- de Boysson-Bardies, B., Sagart, L., & Durand, C. (1984). Discernible differences in the babbling of infants according to target language. *Journal of Child Language*, 11(1), 1-15.

- de Boysson-Bardies, B., Sagart, L., Halle, P., & Durand, C. (1986). Acoustic investigations of cross-linguistic variability in babbling. In B. Lindblom & R. Zetterstrom (Eds.), *Precursors of early speech* (pp. 113-126). New York: Stockton Press.
- de Boysson-Bardies, B., & Vihman, M. (1991). Adaptation to Language: Evidence from babbling and first words in four Languages. *Language*, 67(2), 297-319. doi: citeulike-article-id:7501879
- Dinnsen, D. A. (1992). Variation in developing and fully developed phonetic inventories. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 191-210). Timonium, Maryland: York Press
- Dinnsen, D. A., Chin, S. B., Elbert, M., & Powell, T. W. (1990). Some constraints on functionally disordered phonologies: Phonetic inventories and phonotactics. *Journal of Speech and Hearing Research*, 33(1), 28-37.
- Eilers, R. E., & Oller, D. K. (1994). Infant vocalizations and the early diagnosis of severe hearing impairment. *The Journal of Pediatrics*, 124(2), 199-203. doi: S0022347694003288 [pii]
- Eilers, R. E., Wilson, W. R., & Moore, J. M. (1977). Developmental changes in speech discrimination in infants. *Journal of Speech and Hearing Research*, 20(4), 766-780.
- Eimas, P. (1974). Auditory and linguistic processing of cues for place of articulation by infants. *Perception and Psychophysics*, 16(3), 513-521. doi: 10.3758/bf03198580
- Eimas, P., Siqueland, E. R., Jusczyk, P. W., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171(968), 303-306.
- Engstrand, O., Williams, K., & Lacerda, F. (2003). Does babbling sound native? Listener responses to vocalizations produced by Swedish and American 12- and 18-month-olds. *Phonetica*, 60(1), 17-44. doi: 10.1159/000070452



- Escudero, P., & Polka, L. (2003). *A cross-language study of vowel categorization and vowel acoustics: Canadian English versus Canadian French*. Paper presented at the 15th International Conference of the Phonetic Sciences, Barcelona, Spain.
- Farwanah, S. (2007). Epenthesis. In K. Versteegh (Ed.), *Encyclopedia of Arabic language and linguistics* (Vol. 2, pp. 35-39). Leiden; Boston: Brill.
- Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language*, 51, 419–439.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, 8, 181-195.
- Field, A. P. (2009). *Discovering statistics using SPSS: (And sex and drugs and rock 'n' roll)* (3rd ed.). London: SAGE.
- Fletcher, S. G., & Daly, D. A. (1974). Sublingual dimensions in infants and young children. *Arch Otolaryngol*, 99(4), 292-296. doi: 10.1001/archotol.1974.00780030302014
- Fougeron, C., & Smith, C. L. (1999). French. In International Phonetic Association (Ed.), *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet* (pp. 78-81). Cambridge, UK: Cambridge University Press.
- Gay, L. R., Mills, G. E., & Airasian, P. W. (2012). *Educational research: Competencies for analysis and applications* (10th ed.). Boston, MA: Pearson.
- Gervain, J., & Werker, J. F. (2008). How infant speech perception contributes to language acquisition. *Language and Linguistics Compass*, 2(6), 1149-1170.
- Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences*, 100(13), 8030-8035. doi: 10.1073/pnas.1332441100

- Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, 19(5), 515-523. doi: 10.1111/j.1467-9280.2008.02117.x
- Goldstein, M. H., & Schwade, J. A. (2009). From birds to words: Perception of structure in social interactions guides vocal development and language learning. In M. S. Blumberg, J. H. Freeman & S. R. Robinson (Eds.), *The oxford handbook of developmental and comparative neuroscience* (pp. 708–729). Oxford: Oxford University Press.
- Green, J. R., Moore, C. A., Higashikawa, M., & Steeve, R. W. (2000). The physiologic development of speech motor control: Lip and jaw coordination. *Journal of Speech, Language, and Hearing Research*, 43(1), 239-255.
- Green, J. R., Moore, C. A., & Reilly, K. J. (2002). The sequential development of jaw and lip control for speech. *Journal of Speech, Language, and Hearing Research*, 45(1), 66-79.
- Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders*, 39(5), 350-365. doi: 10.1016/j.jcomdis.2006.06.013
- Hallé, P. A., de Boysson-Bardies, B., & Vihman, M. M. (1991). Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by Japanese and French infants. *Language and Speech*, 34(4), 299-318. doi: 10.1177/002383099103400401
- Hepper, P. G., & Shahidullah, B. S. (1994). Development of fetal hearing. *Archives of Disease in Childhood*, 71(2), F81-87.
- Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97(5 Pt 1), 3099-3111.
- Holes, C. (1990). *Gulf Arabic*. London; New York: Routledge.

- Holes, C. (2004). *Modern Arabic: Structures, functions, and varieties* (Rev. ed.). Washington, D.C.: Georgetown University Press.
- Houston, D. M. (2011). Infant speech perception. In R. Seewald & A. M. Tharpe (Eds.), *Comprehensive Handbook of Pediatric Audiology* (pp. 47-62). San Diego, CA: Plural Publishing, Inc.
- Huthaily, K. (2003). *Contrastive phonological analysis of Arabic and English*. Thesis (MA), University of Montana 2003.
- Ingham, B. (2008). Najdi Arabic. In K. Versteegh (Ed.), *Encyclopedia of Arabic language and linguistics* (Vol. 3, pp. 326-334). Leiden; Boston: Brill.
- Ingram, D. (1989). *First language acquisition: Method, description, and explanation*. Cambridge [England]; New York: Cambridge University Press.
- Ingram, D. (2008). Cross-linguistic phonological acquisition. In M. J. Ball (Ed.), *The Handbook of Clinical Linguistics* (pp. 626-640). Malden, MA; Oxford: Blackwell Pub.
- Ireton, H. (2005). *Child Development Inventory*. Minneapolis, MN: Behavior Science Systems Inc.
- Jakobson, R. (1941/1968). *Child language. Aphasia and phonological universals*. The Hague, Paris,: Mouton.
- Jusczyk, P. W., Copan, H., & Thompson, E. (1978). Perception by 2-month-old infants of glide contrasts in multisyllabic utterances. *Perception and Psychophysics*, 24(6), 515-520. doi: 10.3758/bf03198777
- Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in english-learning infants. *Cognitive Psychology*, 39(3-4), 159-207. doi: 10.1006/cogp.1999.0716

- 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, Maryland: York Press
- Kent, R. D., & Bauer, H. R. (1985). Vocalizations of one-year-olds. *Journal of Child Language*, 12(3), 491-526.
- Kent, R. D., & Murray, A. D. (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *Journal of the Acoustical Society of America*, 72(2), 353-365.
- Kent, R. D., & Read, C. (2002). *The acoustic analysis of speech* (2nd ed.). Albany, NY: Singular/Thompson.
- Kisilevsky, B. S., Hains, S. M., Lee, K., Xie, X., Huang, H., Ye, H. H., . . . Wang, Z. (2003). Effects of experience on fetal voice recognition. *Psychological science*, 14(3), 220-224.
- Kisilevsky, B. S., & Low, J. A. (1998). Human fetal behavior: 100 Years of Study. *Developmental Review*, 18(1), 1-29. doi: DOI: 10.1006/drev.1998.0452
- Kisilevsky, B. S., Pang, L., & Hains, S. M. (2000). Maturation of human fetal responses to airborne sound in low- and high-risk fetuses. *Early Human Development*, 58(3), 179-195. doi: S037837820000075X [pii]
- Klein, J. O. (2000). The burden of otitis media. *Vaccine*, 19 Suppl 1, S2-8.
- Koopmans-van Beinum, F. J., & van der Stelt, J. M. (1986). Early stages in the development of speech movements. In B. Lindblom & R. Zetterstrom (Eds.), *Precursors of early speech* (pp. 37-50). New York: Stockton.
- Kuhl, P. K. (1979). Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America*, 66(6), 1668-1679.

- Kuhl, P. K. (1994). Learning and representation in speech and language. *Current Opinion in Neurobiology*, 4(6), 812-822.
- Kuhl, P. K. (2003). Human speech and birdsong: Communication and the social brain. *Proceedings of the National Academy of Sciences*, 100(17), 9645-9646. doi: 10.1073/pnas.1733998100
- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience* 5(11), 831-843. doi: nrn1533 [pii]10.1038/nrn1533
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., & et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277(5326), 684-686.
- Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, 218(4577), 1138-1141. doi: 10.1126/science.7146899
- Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America*, 100(4 Pt 1), 2425-2438.
- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9(2), F13-F21. doi: 10.1111/j.1467-7687.2006.00468.x
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & et al. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255(5044), 606-608.
- Ladefoged, P. (1999). American English. In International Phonetic Association (Ed.), *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet* (pp. 41-44). Cambridge, UK: Cambridge University Press.

- Ladefoged, P. (2004). *Vowels and consonants* (2nd ed.). Oxford: Blackwell.
- Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford, UK ; Cambridge, Mass.: Blackwell.
- Lavigne-Rebillard, M., & Pujol, R. (1987). Surface aspects of the developing human organ of corti. *Acta Oto-laryngologica*, 104(s436), 43-50. doi: doi:10.3109/00016488709124975
- Lee, S. A. S., Davis, B., & Macneilage, P. (2008). Segmental properties of input to infants: A study of Korean. *Journal of Child Language*, 35(03), 591-617. doi: doi:10.1017/S0305000908008684
- Lee, S. A. S., Davis, B., & Macneilage, P. (2010). Universal production patterns and ambient language influences in babbling: A cross-linguistic study of Korean- and English-learning infants. *Journal of Child Language*, 37(02), 293-318. doi: doi:10.1017/S0305000909009532
- Levitt, A. G., Jusczyk, P. W., Murray, J., & Carden, G. (1988). Context effects in two-month-old infants' perception of labiodental/interdental fricative contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 14(3), 361-368. doi: 10.1037/0096-1523.14.3.361
- Levitt, A. G., & Utman, J. G. (1992). From babbling towards the sound systems of English and French: A longitudinal two-case study. *Journal of Child Language*, 19(01), 19-49. doi: doi:10.1017/S0305000900013611
- 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.
- Lewis, M. P. (2009). Ethnologue: Languages of the world Retrieved from Online version: <http://www.ethnologue.com>

- Lieberman, D. E., McCarthy, R. C., Hiiemae, K. M., & Palmer, J. B. (2001). Ontogeny of postnatal hyoid and larynx descent in humans. *Archives of oral biology*, 46(2), 117-128.
- Liu, H. M., Tsao, F. M., & Kuhl, P. K. (2005). The effect of reduced vowel working space on speech intelligibility in Mandarin-speaking young adults with cerebral palsy. *Journal of the Acoustical Society of America*, 117(6), 3879-3889. doi: 10.1121/1.1898623
- Locke, J. L. (1983). *Phonological acquisition and change*. New York: Academic Press.
- Macken, M. A. (1995). Phonological acquisition. In J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 671-696). Cambridge, Mass., USA: Blackwell.
- MacNeilage, P. F., & Davis, B. L. (2000). On the origin of internal structure of word forms. *Science*, 288(5465), 527-531. doi: 8452 [pii]
- MacNeilage, P. F., & Davis, B. L. (2001). Motor mechanisms in speech ontogeny: Phylogenetic, neurobiological and linguistic implications. *Current Opinion in Neurobiology*, 11(6), 696-700. doi: S0959-4388(01)00271-9 [pii]
- Malécot, A. (1974). Frequency of occurrence of French phonemes and consonant clusters. *Phonetica*, 29(3), 158-170.
- Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science*, 11(1), 122-134. doi: DESC653 [pii] 10.1111/j.1467-7687.2007.00653.x
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82(3), B101-111. doi: S0010027701001573 [pii]
- Menard, L., Schwartz, J.-L., & Boe, L.-J. (2004). Role of vocal tract morphology in speech development: Perceptual targets and sensorimotor maps for synthesized

- French vowels from birth to adulthood. *Journal of Speech, Language, and Hearing Research*, 47(5), 1059-1080. doi: 10.1044/1092-4388(2004/079)
- Menn, L. (1983). Development of articulatory, phonetic, and phonological capabilities. In B. Butterworth (Ed.), *Language production* (Vol. 2, pp. 1–49). London: Academic Press.
- Mines, M. A., Hanson, B. F., & Shoup, J. E. (1978). Frequency of occurrence of phonemes in conversational English. *Language and Speech*, 21(3), 221-241. doi: 10.1177/002383097802100302
- Mitchell, T. F. (1993). *Pronouncing Arabic* (Vol. 2). Oxford: Clarendon Press.
- Molemans, I., Van Den Berg, R., Van Severen, L., & Gillis, S. (2012). How to measure the onset of babbling reliably? *Journal of Child Language*, 39(03), 523-552. doi: doi:10.1017/S0305000911000171
- Molnar, M., & Polka, L. (2004). *Vowel preferences in infancy: The role of early language experience*. Paper presented at the Boston Child Language Conference, Boston, MA.
- Moore, C. A. (2004). Physiological development of speech production. In B. Maassen (Ed.), *Speech motor control in normal and disordered speech* (pp. 191–209). Oxford ; New York: Oxford University Press.
- Moore, C. A., Caulfield, T. J., & Green, J. R. (2001). Relative kinematics of the rib cage and abdomen during speech and nonspeech behaviors of 15-month-old children. *Journal of Speech and Hearing Research*, 44(1), 80-94. doi: 10.1044/1092-4388(2001/008)
- Moore, C. A., & Ruark, J. L. (1996). Does speech emerge from earlier appearing oral motor behaviors? *Journal of Speech and Hearing Research*, 39(5), 1034-1047.
- Moore, C. A., Smith, A., & Ringel, R. L. (1988). Task-specific organization of activity in human jaw muscles. *Journal of Speech and Hearing Research*, 31(4), 670-680.



- Moore, J. K. (2002). Maturation of human auditory cortex: Implications for speech perception. *Annals of Otolology Rhinology and Laryngology Supplement*, 189, 7-10.
- Moore, J. K., & Linthicum, F. H., Jr. (2007). The human auditory system: A timeline of development. *International Journal of Audiology* 46(9), 460-478. doi: 781869369 [pii]  
10.1080/14992020701383019
- Moore, J. K., Perazzo, L. M., & Braun, A. (1995). Time course of axonal myelination in the human brainstem auditory pathway. *Hearing Research*, 87(1-2), 21-31.
- Nittrouer, S. (2001). Challenging the notion of innate phonetic boundaries. *Journal of the Acoustical Society of America*, 110(3 Pt 1), 1598-1605.
- Novitski, N., Huotilainen, M., Tervaniemi, M., Näätänen, R., & Fellman, V. (2007). Neonatal frequency discrimination in 250-4000-Hz range: Electrophysiological evidence. *Clinical Neurophysiology*, 118(2), 412-419. doi: DOI: 10.1016/j.clinph.2006.10.008
- Novitski, N., Tervaniemi, M., Huotilainen, M., & Näätänen, R. (2004). Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures. *Cognitive Brain Research*, 20(1), 26-36. doi: DOI: 10.1016/j.cogbrainres.2003.12.011
- 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* (pp. 93-112). New York: Academic Press.
- Oller, D. K. (1986). Metaphonology and infant vocalizations. In B. Lindblom & R. Zetterstrom (Eds.), *Precursors of early speech* (pp. 21-35). New York: Stockton Press.

- Oller, D. K. (2000). *The emergence of the speech capacity*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Oller, D. K., & Eilers, R. E. (1982). Similarity of babbling in Spanish- and English-learning babies. *Journal of Child Language*, 9(3), 565-577.
- Oller, D. K., & Eilers, R. E. (1988). The role of audition in infant babbling. *Child Development*, 59(2), 441-449.
- Oller, D. K., & Eilers, R. E. (1998). Interpretive and methodological difficulties in evaluating babbling drift. *Revue Parole*, 7/8, 147-164.
- Oller, D. K., Eilers, R. E., Bull, D. H., & Carney, A. E. (1985). Prespeech vocalizations of a deaf Infant: A comparison with normal metaphonological development. *Journal of Speech and Hearing Research*, 28(1), 47-63.
- Oller, D. K., & Lynch, M. P. (1992). Infant vocalizations and innovations in infraphonology: Toward a broader theory of vocal development and disorders. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (Eds.), *Phonological Development: Models, Research, Implications* (pp. 509–536). Timonium, Maryland: York Press
- Olsho, L. W., Koch, E. G., Carter, E. A., Halpin, C. F., & Spetner, N. B. (1988). Pure-tone sensitivity of human infants. *Journal of the Acoustical Society of America*, 84(4), 1316-1324.
- Omar, M. K. (1973). *The acquisition of Egyptian Arabic as a native language*. The Hague; Paris: Mouton.
- Patterson, M. L., & Werker, J. F. (1999). Matching phonetic information in lips and voice is robust in 4.5-month-old infants. *Infant Behavior and Development*, 22(2), 237-247. doi: 10.1016/s0163-6383(99)00003-x

- Patterson, M. L., & Werker, J. F. (2003). Two-month-old infants match phonetic information in lips and voice. *Developmental Science*, 6(2), 191-196. doi: 10.1111/1467-7687.00271
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Belmont, CA, etc.: Wadsworth Thomson Learning.
- Peterson-Falzone, S. J., & Monoson, P. K. (2003). *Vocal tract anatomy and functional consequences in cleft lip/palate and syndromes of craniosynostosis*: John Wiley & Sons, Inc.
- Polka, L., & Bohn, O.-S. (2003). Asymmetries in vowel perception. *Speech Communication*, 41(1), 221-231. doi: 10.1016/s0167-6393(02)00105-x
- Polka, L., & Bohn, O.-S. (2011). Natural referent vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, 39(4), 467-478. doi: 10.1016/j.wocn.2010.08.007
- Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d /-- / [edh] / perception: Evidence for a new developmental pattern. *Journal of the Acoustical Society of America*, 109(5), 2190-2201.
- Polka, L., Rvachew, S., & Mattock, K. (2007). Experiential influences on speech perception and production in infancy. In E. Hoff & M. Shatz (Eds.), *Handbook of Language Development* (Vol. 153-172). Malden, MA: Oxford: Blackwell Publishing.
- Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 20(2), 421-435.
- Pujol, R., & Lavigne-Rebillard, M. (1992). Development of neurosensory structures in the human cochlea. *Acta Otolaryngol*, 112(2), 259-264.

- Pye, C., Ingram, D., & List, H. (1987). A comparison of initial consonant acquisition in English and Quiche. In K. E. Nelson & A. Van Kleeck (Eds.), *Children's Language* (pp. 175-190). Hillsdale, NJ: Lawrence Erlbaum.
- Rance, G., & Tomlin, D. (2006). Maturation of auditory steady-state responses in normal babies. *Ear and Hearing*, 27(1), 20-29. doi: 10.1097/01.aud.0000191944.03797.5a
- Reetz, H., & Jongman, A. (2009). *Phonetics: Transcription, production, acoustics, and perception*. Chichester: Blackwell.
- Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science*, 8(2), 162-172. doi: 10.1111/j.1467-7687.2005.00403.x
- Rommel, N., Bellon, E., Hermans, R., Smet, M., De Meyer, A.-M., Feenstra, L., . . . Veereman-Wauters, G. (2003). Development of the orohypopharyngeal cavity in normal infants and young children. *The Cleft Palate-Craniofacial Journal*, 40(6), 606-611. doi: doi:10.1597/1545-1569(2003)040<0606:DOTOCI>2.0.CO;2
- Ruark, J. L., & Moore, C. A. (1997). Coordination of lip muscle activity by 2-year-old children during speech and nonspeech tasks. *Journal of Speech, Language, and Hearing Research*, 40(6), 1373-1385.
- Rvachew, S., Alhaidary, A., Mattock, K., & Polka, L. (2008). Emergence of the corner vowels in the babble produced by infants exposed to Canadian English or Canadian French. *Journal of Phonetics*, 36(4), 564-577.
- Rvachew, S., & Brosseau-Lapr , F. (2010). Speech perception intervention. In S. M. L. Williams & R. McCauley (Eds.), *Interventions For Speech Sound Disorders in Children*. (pp. 295-314). Baltimore, Maryland: Paul Brookes Publishing Co.
- Rvachew, S., Creighton, D., Feldman, N., & Sauve, R. (2002). Acoustic-phonetic description of infant speech samples: Coding reliability and related methodological issues. *Acoustics Research Letters Online*, 3(1), 24-28.

- Rvachew, S., Mattock, K., Polka, L., & Menard, L. (2006). Developmental and cross-linguistic variation in the infant vowel space: The case of Canadian English and Canadian French. *Journal of the Acoustical Society of America*, 120(4), 2250-2259.
- Rvachew, S., Slawinski, E., Williams, M., & Green, C. (1996). Formant frequencies of vowels produced by infants with and without early onset otitis media. *Canadian Acoustics*, 24, 19-28.
- Rvachew, S., Slawinski, E. B., Williams, M., & Green, C. L. (1999). The impact of early onset otitis media on babbling and early language development. *Journal of the Acoustical Society of America*, 105(1), 467-475.
- Saffran, J. R., Werker, J. F., & Werner, L. A. (2006). The infant's auditory world: Hearing, speech, and the beginnings of language. In D. Kuhn & R. S. Siegler (Eds.), *Cognition, perception, and language* (6th ed., pp. 58–108). Hoboken, N.J.: John Wiley & Sons.
- Sander, E. K. (1972). When are speech sounds learned? *Journal of Speech & Hearing Disorders*, 37(1), 55-63.
- Shahin, K. (1996). Assessing pharyngeal place in Palestinian Arabic. In M. Eid & D. Parkinson (Eds.), *Perspectives on Arabic linguistics IX: papers from the Ninth annual symposium on Arabic linguistics*. (pp. 131-149). Amsterdam ; Philadelphia: John Benjamins.
- Siebert, J. R. (1985). A morphometric study of normal and abnormal fetal to childhood tongue size. *Archives of oral biology*, 30(5), 433-440. doi: 10.1016/0003-9969(85)90072-x
- Smith, N. A., Trainor, L. J., & Shore, D. I. (2006). The development of temporal resolution: Between-channel gap detection in infants and adults. *Journal of Speech, Language, and Hearing Research*, 49(5), 1104-1113. doi: 10.1044/1092-4388(2006/079)

- Sperber, G. H. (2001). *Craniofacial development*. Hamilton, Ont.; London: B.C. Decker.
- Spetner, N. B., & Olsho, L. W. (1990). Auditory frequency resolution in human infancy. *Child Development*, 61(3), 632-652.
- 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, pp. 73-92). New York: Academic Press.
- Steeve, R. W., Moore, C. A., Green, J. R., Reilly, K. J., & McMurtrey, J. R. (2008). Babbling, chewing, and sucking: Oromandibular coordination at 9 months. *Journal of Speech, Language, and Hearing Research*, 51(6), 1390-1404. doi: 10.1044/1092-4388(2008/07-0046)
- Stevens, S. S., Volkmann, E. B., & Newman, J. (1937). The mel scale equates the magnitude of perceived differences in pitch at different frequencies. *Journal of the Acoustical Society of America*, 8, 185-190.
- Stoel-Gammon, C. (1988). Prelinguistic vocalizations of hearing-impaired and normally hearing subjects: A comparison of consonantal inventories. *Journal of Speech and Hearing Disorders*, 53(3), 302-315.
- Stoel-Gammon, C., & Otomo, K. (1986). Babbling development of hearing-impaired and normally hearing subjects. *Journal of Speech and Hearing Disorders*, 51(1), 33-41.
- Stoel-Gammon, C., & Pollock, K. E. (2008). Vowel development and disorders. In M. J. Ball (Ed.), *The handbook of clinical linguistics* (pp. 525-548). Malden, MA ; Oxford: Blackwell Pub.
- Stokes, S. F., & Surendran, D. (2005). Articulatory complexity, ambient frequency, and functional load as predictors of consonant development in children. *Journal of Speech Language & Hearing Research*, 48(3), 577-591.

- Streeter, L. A. (1976). Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature*, 259(5538), 39-41.
- Sundara, M., Polka, L., & Genesee, F. (2006). Language-experience facilitates discrimination of /d-/ in monolingual and bilingual acquisition of English. *Cognition*, 100(2), 369-388. doi: 10.1016/j.cognition.2005.04.007
- Swoboda, P. J., Kass, J., Morse, P. A., & Leavitt, L. A. (1978). Memory Factors in Vowel Discrimination of Normal and At-Risk Infants. *Child Development*, 49(2), 332-339.
- Tharpe, A. M., & Ashmead, D. H. (2001). A longitudinal investigation of infant auditory sensitivity. *American Journal of Audiology*, 10(2), 104-112.
- 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, N.J.: L. Erlbaum.
- Thelwall, R., & Sa'adeddin, M. A. (1999). Arabic. In International Phonetic Association (Ed.), *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet* (pp. 51-54). Cambridge, UK: Cambridge University Press.
- Thevenin, D., Eilers, R. E., Oller, D. K., & Lavoie, L. (1985). Where's the drift in babbling drift? A cross-linguistic study. *Applied Psycholinguistics*, 6, 3-15.
- Vihman, M. (1996). *Phonological development: The origins of language in the child*. Cambridge, Mass.: Blackwell.
- Vihman, M. (2006). *Phonological templates in early words: A cross-linguistic study*. Paper presented at the The 10th Conference on Laboratory Phonology, Paris, France.

- Vihman, M., & Croft, W. (2007). Phonological development: Toward a "radical" templatic phonology. *Linguistics*, 45(4), 683-725.  
doi: <http://dx.doi.org/10.1515/LING.2007.021>
- Vorperian, H. K., Kent, R. D., Gentry, L. R., & Yandell, B. S. (1999). Magnetic resonance imaging procedures to study the concurrent anatomic development of vocal tract structures: Preliminary results. *International Journal of Pediatric Otorhinolaryngology*, 49(3), 197-206. doi: 10.1016/s0165-5876(99)00208-6
- Vorperian, H. K., Kent, R. D., Lindstrom, M. J., Kalina, C. M., Gentry, L. R., & Yandell, B. S. (2005). Development of vocal tract length during early childhood: A magnetic resonance imaging study. *The Journal of the Acoustical Society of America*, 117(1), 338-350.
- Vorperian, H. K., Wang, S., Chung, M. K., Schimek, E. M., Durtschi, R. B., Kent, R. D., . . . Gentry, L. R. (2009). Anatomic development of the oral and pharyngeal portions of the vocal tract: An imaging study. *The Journal of the Acoustical Society of America*, 125(3), 1666-1678.
- Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: the privileged status of speech for young infants. *Developmental Science*, 7(3), 270-276. doi: 10.1111/j.1467-7687.2004.00345.x
- Watson, J. C. E. (2002). *The phonology and morphology of Arabic*. Oxford ; New York: Oxford University Press.
- Werker, J. F., & Tees, R. C. (2002). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and development*, 25(1), 121-133.
- Werker, J. F., & Yeung, H. H. (2005). Infant speech perception bootstraps word learning. *Trends in Cognitive Sciences*, 9(11), 519-527. doi: 10.1016/j.tics.2005.09.003
- Werner, L. A. (2002). Infant auditory capabilities. *Current Opinion in Otolaryngology and Head and Neck Surgery*, 10(5), 398-402.



- Werner, L. A., Marean, G. C., Halpin, C. F., Spetner, N. B., & Gillenwater, J. M. (1992). Infant auditory temporal acuity: Gap detection. *Child Development*, 63(2), 260-272.
- Wilson, E. M., Green, J. R., Yunusova, Y., & Moore, C. A. (2008). Task specificity in early oral motor development. *Seminars in Speech and Language*, 29(4), 257-266. doi: 10.1055/s-0028-1103389