VISUAL LANGUAGE DISCRIMINATION

by

WHITNEY MARIE WEIKUM

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Neuroscience)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

February, 2008

ABSTRACT

Recognizing and learning one's native language requires knowledge of the phonetic and rhythmical characteristics of the language. Few studies address the rich source of language information available in a speaker's face. Solely visual speech permits language discrimination in adults (Soto-Faraco et al., 2007). This thesis tested infants and adults on their ability to use only information available in a speaker's face to discriminate rhythmically dissimilar languages.

Monolingual English infants discriminated French and English using only visual speech at 4 and 6 months old, but failed this task at 8 months old. To test the role of language experience, bilingual (English/French) 6 and 8-month-old infants were tested and successfully discriminated the languages. An optimal period for sensitivity to visual language information necessary for discriminating languages may exist in early life.

To confirm an optimal period, adults who had acquired English as a second language were tested. If English was learned before age 6 years, adults discriminated English and French, but if English was learned after age 6, adults performed at chance. Experience with visual speech information in early childhood influences adult performance.

To better understand the developmental trajectory of visual language discrimination, visual correlates of phonetic segments and rhythmical information were examined. When clips were manipulated to remove rhythmical information, infants used segmental visual phonetic cues to discriminate languages at 4, but not 8 months old. This suggests that a decline in non-native visual phonetic discrimination (similar to the decline seen for non-native auditory phonetic information; Werker & Tees, 1984), may be impairing language discrimination at 8 months.

Infants as young as newborn use rhythmical auditory information to discriminate languages presented forward, but not backward (Mehler et al., 1988). This thesis showed that both 4 and 8-month-old infants could discriminate French from English when shown reversed language clips. Unlike auditory speech, reversed visual speech must conserve cues that permit language discrimination.

Infants' abilities to distinguish languages using visual speech parallel auditory speech findings, but also diverge to highlight unique characteristics of visual speech. Together, these studies further enrich our understanding of how infants come to recognize and learn their native language(s).

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

Thank you for the inspiration, constant commitment,
unfaltering reinforcement and lighthearted repartee.
With much gratitude to…

Reed Weikum, for leading by example.
Sharon and Mel Weikum for your unconditional love.

DEDICATION

For My Mother
No Matter What

CO-AUTHORSHIP STATEMENT

The identification and design of the research program was developed in discussion with the research supervisor, Janet Werker. The adult and infant studies were designed to complement a line of visual language discrimination research developed by Janet Werker, Núria Sebastián-Gallés, and Salvador Soto-Faraco with the assistance of Athena Vouloumanos and Jordi Navarra.

The stimuli used for the experiments were initially prepared by Jordi Navarra and Athena Vouloumanos, and edited and manipulated by the author for each study in this thesis. Manipulations of the visual speech stimuli were primarily designed by the author, in discussion with the research supervisor Janet Werker and co-author Eric Vatikiotis-Bateson.

The research was performed, data was analyzed, and manuscripts were prepared by the author, with the assistance of the research supervisor Janet Werker. The co-authors listed for each paper have/will assist with final editing of the manuscripts.

CHAPTER I – **Visual Language Discrimination Introduction**

1.1 Introduction

Knowledge of the phonetic and rhythmical regularities characteristic of one's native language(s) is necessary for eventually recognizing, perceiving and producing the language(s). Considerable research has revealed infants' developing sensitivity to the phonetic (e.g. Werker & Tees, 1984) and rhythmical (e.g. Nazzi, Bertoncini & Mehler, 1998) information of their native language. Previous work has focused almost exclusively on information available in the acoustic signal, overlooking the rich source of language information available in the speaker's head and face. From birth, an infant's language experience is dominated by dynamic talking heads that provide not only auditory, but rich visual language information. Exploring how visual speech information alone may contribute to language processing is the focus of the studies described in this thesis. Primarily, the studies presented in this thesis examined whether discrimination of rhythmically dissimilar languages is possible in both infants and adults. Infants and adults were tested to determine when the ability to discriminate languages visually develops, as well as the factors influencing visual language discrimination in both infancy and adulthood. Different ages and language combinations were examined to help determine whether or not periods of sensitivity to visual speech information exist in infancy and further influence adult language discrimination performance. Findings from these studies highlight the visual speech cues that are sufficient for language discrimination in infancy, and contribute to our understanding of how infants and adults may use visual speech information to recognize and learn language(s). Examining developmental trajectories for visual speech perception also contributes to research addressing whether infants are initially prepared to learn language (or languages) using more than one modality.

1.2 Role of Visual Information in Speech Perception Development

Infants grow up in noisy, dynamic, rich language environments where speech is not always presented in clear, quiet surroundings with only one speaker present. Over half the world's population is bilingual (Brutt-Griffler & Varghese, 2004), contributing to the complexity of the speech signal that infants may experience. When learning language, infants must be able to accurately attend to a single speaker in noisy, multi-speaker and/or multi-language environments. Understanding how infants come to select and preferentially attend to speakers of their native language in noisy or ambiguous environments may hold clues to the methods used by infants to isolate the speech stream, a necessary step when acquiring language. Determining whether infants can rely on solely visual speech information to identify speakers of their native language helps determine if visual speech is a method by which speakers can be classified. Secondly, examining how infants and adults perform on visual language discrimination tasks at different ages, and with exposure to different languages, helps determine the developmental trajectory of visual language perception. Finally, experiments investigating the facial cues that infants use to discriminate languages help determine the specific visual speech properties that infants are sensitive to and use to process language information.

1.2.1 Developmental Trajectory for Visual Speech Information

By birth infants appear to be well prepared to learn speech auditorily, but it is unclear how visual speech perception may develop in infancy. It could be that infants, through experience with visual information, develop the ability to recognize visual language information by hearing and seeing speech at the same time. If this were the case, it might be expected that infants may not discriminate languages visually at a young age, or discriminate poorly and then

become measurably better over some period of time (for instance the first year of life). If however, visual speech information patterns similarly to auditory speech information, infants may come prepared to discriminate languages visually, and then maintain or lose the ability to discriminate visual language information, depending on whether the relevant information is available in their language-learning environment.

Examining the developmental trajectory of visual speech perception enables a clearer understanding of how the speech signal may be integrated multi-modally. Some theorists propose that speech may have evolved as a multimodal signal (e.g. MacNeilage, 1998). Multi-modal vs. uni-modal presentation has been shown to enhance speed and accuracy on tasks using auditory and/or visual information (e.g. Vroomen & de Gelder, 2000). Behavioural studies show that the integration of auditory and visual information appears to occur prior to phonetic categorization (e.g. Summerfield, 1987; Green, 1998) and neurophysiological evidence suggests that areas of the brain, thought to be for the exclusive use of auditory speech information, are activated by solely visual speech information as well (e.g. Calvert et al. 1997). These conclusions, however, are not without their critics. For example, some studies have revealed activation of secondary auditory cortex related to language processing, but failed to find activation of primary auditory cortex to solely visual speech information (e.g. Bernstein et al., 2002) and Massaro's (1987) research supports a view whereby auditory and visual information are processed separately and then mapped onto a phonetic prototype.

The primacy or pre-categorization shown by behavioural studies that integrate both auditory and visual speech information, coupled with the overlapping activation of speech areas to both auditory and visual stimuli, suggests that speech may be modality-neutral. Indeed, several theorists have suggested that speech perception may stem from modality independent

information (e.g. Gibson, 1966; Summerfield, 1987). Rosenblum (2005) argues that cross-modal integration is not something that occurs within the speaker, but is an inherent property of the speech information. For instance, Rosenblum, Miller and Sanchez (2007) have shown that allowing a person to become familiar with only a person's visual speech information will later facilitate recovery of the person's auditory speech signal. This is taken as evidence that modality neutral information may be transmitted through familiarity with a person's visual speech.

If speech is in fact, modality neutral, another supporting finding would be that visual speech perception, for instance language discrimination, would pattern similarly to auditory speech development. Experiments designed to test robust findings from the auditory speech literature using visual speech information would help to decipher whether visual speech information contributes similar language information to the speech percept, or has a different developmental trajectory, thereby suggesting the dissemination of different speech information during development. For instance, over the first year of life, infants' sensitivity to non-native phonetic (Werker & Tees, 1984) and facial (Pascalis et al., 2005) information declines. If visual language information shares the same foundation, the ability to tell languages apart visually should also occur early in infancy, and a decline in sensitivity to native vs. non-native speech information should occur over the first year of life.

The studies described in this thesis tested whether there is a broad based sensitivity to visual language information and subsequent developmental decline that is similar to that reported for the auditory phonetic and facial information described above. The first set of studies described in this thesis examined whether infants can tell languages apart visually, and if there are differences in performance across the first year of life. To determine if the exposure to visual language information in infancy has an impact on adult speech perception, studies were then

conducted to examine whether languages must be acquired in infancy in order for adults to be able to use visual speech information to discriminate languages. If there is some sensitivity and (re)organization[1] of visual language information during infancy or early childhood that is necessary for later visual language discrimination, it was expected that adults who learned a second language late in childhood may be impaired when discriminating their second language from an unfamiliar language. And finally, the cues that infants are sensitive to in the visual speech signal were examined to determine the visual information sufficient for language discrimination. The cues examined include dynamic configural information from the articulators and corresponding head and face (visual phonetic cues) as well as patterned timing information conveyed by the articulators and corresponding head and face movements (rhythmical cues). Investigating whether the visual speech information necessary for language discrimination corresponds to the auditory information necessary for differentiating languages, helped determine the specific visual speech cues necessary for language discrimination. It also helped to determine the visual speech information that might experience a decline in sensitivity following an optimal period of sensitivity for visual language information that may exist in early infancy and childhood.

1.2.2 Adult Visual Speech Perception

Visual speech perception studies with adults have robustly shown that when the face is visible, it can dramatically improve speech intelligibility in noisy situations (Gagne, Masterson, & Munhall, 1994; Helfter, 1997; Sumby & Pollack, 1954; Summerfield, 1979). Even clear speech, if heavy in semantic content or strongly accented, has been shown to benefit from the

---

[1] The "(re)" attached to "organization" is placed in brackets to clarify that this "reorganization" is part of an ongoing organization that has started before and will continue following any (re)organization related to perceptual processes during the infancy period.

addition of visual speech information (Reisberg, McLean, & Goldfield, 1987). Furthermore, mismatching visual and acoustic information can change the speech percept (for example, a visual /ga/ paired with an auditory /ba/ leads to the percept /da/; McGurk & MacDonald, 1976). Functional magnetic resonance imaging (fMRI) studies have even shown how areas related to language are activated by visual speech information. Significant activations in Broca's area occurr in response to a face silently reciting words compared to a still face (Olson, Gatenby & Gore, 2002). Calvert and colleagues (1997) showed that silent visual lipreading of the numbers 1-9 activates areas that include both primary auditory and auditory association cortices, whereas facial gurning (mouth movements unrelated to speech) fails to activate the auditory regions, and instead activates areas related to attention-demanding tasks. Other studies examining the effects of lipreading in the absence of scanner noise (MacSweeney et al., 2000) and pure tone vs. lipreading activation (Bernstein et al., 2002) have also revealed evidence for the activation of secondary auditory cortex associated with processing language, although they failed to find activation of primary auditory cortex.

Information conveyed by the lips and mouth that correspond to the sound repertoire characteristic of a language have traditionally been treated as visual representations of phonemes or visemes. However, dynamic information in the rest of the face (e.g. Thomas & Jordan, 2004; Vatikiotis-Bateson, Eigsti, Yano, & Munhall, 1998) and even head motions (e.g. Munhall, Jones, Callan, Kuratate, & Vatikiotis-Bateson, 2004) contribute to the intelligibility of speech. Research has also demonstrated the surprising level of detail available in visual speech, in that it can convey information as specific as the voice of the speaker (Kamachi, Hill, Lander & Vatikiotis-Bateson, 2003). Even though adults are notoriously poor at lipreading (e.g. Bernstein, Demorest, & Tucker, 2000), they are capable of discriminating languages using only visual

speech information available in a speaker's head and face (Soto-Faraco et al., 2007). Soto-Faraco and colleagues (2007) had adults watch the face of a bilingual speaker silently reciting sentences in both Spanish and Catalan. After watching two consecutive sentences, the adults indicated whether they thought the sentences were from the "same" or "different" languages. Adults who were not familiar with either language performed at chance, but adults who were familiar with one or both languages performed significantly better than chance. The adults' proficiency at recognizing words from the sentences was tested, and based on the number of words correctly identified, it was statistically unlikely that the adults used word recognition through lip-reading to perform the task. This suggests that even when the adults do not understand the words being said, they are able to extract information sufficient for telling languages apart. Given that word recognition is not the basis for visual language discrimination, it is possible that infants, before they understand the words of the language, may be capable of using visual speech cues to tell apart the languages.

1.2.3 Infant Visual Speech Perception

Speech directed to infants is richly multimodal (e.g. Sullivan & Horowitz, 1983), so visual speech information may provide an additional set of cues that infants can use to discriminate and ultimately learn languages. Indeed, studies with blind individuals suggest that visual speech information influences language acquisition. Infants who are born blind or significantly visually impaired show delays in language acquisition (e.g. longer babbling period, delays in first words) and phonological errors that are inconsistent with the errors shown by typically sighted children (Mills, 1983). In the Mills research, the spontaneous utterances of blind and typically sighted children, aged 2-3 years, were examined. It was found that a blind

child makes proportionately more phonemic substitutions between visually dissimilar speech categories, rather than within visually similar speech categories (Mills, 1983). This suggests that sighted children make use of visual information to learn phonemic categories more rapidly and efficiently. Thus, it is not unreasonable to expect that sighted infants may be responsive to, and influenced by, visual speech cues. It might even be the case that an integrated speech percept is more efficient for language learning.

Infants have been shown to be sensitive to both segmental visual phonetic and rhythmical speech information. At 2-4 months, infants can match an auditorily heard vowel with the appropriate face mouthing the vowel (Kuhl & Meltzoff, 1984; Patterson & Werker, 2003). When the auditory and visual information are in conflict (e.g. visual /ba/ and auditory /ga/), 4-5 month old infants' "heard" percept may also be influenced by what they see (Burnham & Dodd, 2004; Desjardins & Werker, 2004; Rosenblum, Schmuckler, & Johnson, 1997). By 10-14 weeks infants prefer watching faces of people reciting nursery rhymes when the speech sounds and lip movements are in synchrony (Dodd, 1979). Additionally, infants presented with two faces talking side-by-side prefer watching the face that corresponds to the auditorily heard passage at 4.5 months, but not 2.5 months (Dodd & Burnham, 1988). However, this latter result only occurs on trials where the infant is matching their native language, making it unclear whether infants are matching the audio-visual synchrony, or simply preferring to watch their native language.

Infants show sensitivity to visual rhythmical information that is necessary for parsing the speech stream. Hollich, Newman, and Jusczyk (2005) showed that infants (7.5 months) are sensitive to rhythmical waveform information because they are capable of using a visual wave to help parse the speaker's voice that matches the waveform from a distracting voice. The infants

are able to use a visual waveform representing the timing and amplitude of the speech signal to attend to a speaker's voice that they are normally incapable of parsing through the noise. Because there is a strong degree of correlation between the auditory signal amplitude and the dynamic visual information expressed by multiple sites on the face and head (Yehia, Rubin & Vatikiotis-Bateson, 1998), it may be the case that strictly visual speech information is sufficient for conveying rhythmical language information necessary for distinguishing between languages.

Thus, there is considerable evidence that infants do attend to and use visual information when processing speech in multi-speaker and multi-lingual environments. The infant studies in this thesis will further advance our understanding of the role of visual information by determining the circumstances under which infants are capable of visually discriminating languages.

1.3 Infant Speech Perception

Precisely how visual speech information is perceived and develops with regard to auditory speech has been rarely examined. Speech perception studies have shown that at birth, infants prefer the sound of human speech over similarly complex non-speech (e.g. Vouloumanos & Werker, 2007) and can tell the difference between rhythmically unfamiliar languages (e.g. Mehler et al., 1988). Young infants can also discriminate nearly all the speech sounds on which they have been tested, including sounds not used in the native language (Streeter, 1976; Trehub, 1976; Werker, Gilbert, Humphrey, & Tees, 1981). Over the first year of life their ability to discriminate the sound contrasts from their native language is maintained or improves (Kuhl et al., 2006), and boundaries are realigned (Burns, Yoshida, Hill, & Werker, 2007), whereas discrimination performance on non-native speech contrasts declines [Best, McRoberts, LaFleur

& Silver-Isenstadt, 1995; Mattock & Burnham, 2006; Werker & Tees, 1984; for a review see Werker & Tees, 1999; see Best, McRoberts, & Sithole, 1988 for an exception].

1.4 Perceptual Development Trajectories

Developmental studies examining the perceptual abilities of infants have encountered robust and dramatic changes in perception over the first year of life. These changes in sensitivity have been found across a number of domains, including language, face and music perception. For instance, Werker and Tees (1984) showed that infants' sensitivity to consonant contrasts that are not part of their native language significantly reduces by 10-12 months of age. A similar effect occurs a bit earlier for vowels (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Polka & Werker, 1994) and recently, Mattock and Burnham (2006) found that this effect even occurs for lexical tones. Infants hearing tone languages continue to discriminate lexical tones at 9 months, but infants who do not hear tone languages experience a decline in sensitivity.

The decline in phonetic perception sensitivity can be reversed if infants are exposed to non-native contrasts during the period when sensitivity normally declines (Kuhl, Tsao, & Liu, 2003). Furthermore, continued exposure to contrasts from an infant's native language appears to improve discrimination (Kuhl et al., 2006).

With regard to face perception, Pascalis, de Haan and Nelson (2002) showed that infants discriminate both monkey and human faces at 6 months, but 9-month-olds and adults no longer discriminate monkey faces. However, exposure to monkey faces between 6 and 9 months allows the infants to maintain their discrimination of the monkey faces (Pascalis et al., 2005). Lewkowicz and Ghazanfar (2006) have shown that even the ability to pair non-species face and vocalization matchings declines over the first year of life.

Perceptual changes across the first year of life are not limited to auditory speech and face perception. Visual speech perception in the form of sign language shows similar changes. Hearing infants, at 4 months, differentiate hand shapes corresponding to different categorical boundaries characteristic of American Sign Language, but infants at 14 months no longer discriminate these hand shapes according to their categorical boundaries (Baker, Golinkoff & Petitto, 2006). Research with musical rhythms from other cultures (Hannon & Trehub, 2005) similarly shows that young infants can discriminate different cultural rhythms, but this ability also declines by adulthood. These studies demonstrate how infancy is an important time for the (re)organization of perceptual processes.

1.4.1 Optimal Periods In Speech Perception

A considerable amount of perceptual (re)organization occurs in infancy. Whether or not the perceptual foundations laid in early infancy and childhood have a lasting impact on perceptual sensitivities is of pertinent interest. Periods of sensitivity related to the development of both visual and auditory processes have been found across a number of organisms, and research has identified some of the mechanisms responsible for the opening and closing of these periods of sensitivity (for a review see Hensch, 2004). With regard to language processes, Lenneberg (1967) proposed that a critical period exists, starting from birth and lasting until the onset of puberty. Indeed, there are a number of studies suggesting that periods exist in early infancy and childhood for which the foundation for certain language abilities are established (for a review see Werker & Tees, 2005). Tees (2001) has suggested that the term "optimal period" be used to describe periods in development when neural and behavioural functioning is particularly sensitive to certain environmental input.

Studies examining phonological development have found ages beyond which it is more difficult, or nearly impossible to obtain native-like proficiency on some language pronunciation and perception tasks. For instance, in order to acquire an accent free language, it appears that the language needs to be learned before age 6-8 (for a review see Piske, MacKay, & Flege, 2001). Language production of consonants also differs in second language learners depending on the age learned (Flege, 1991). However, some differences have been found on more difficult phonological tasks, even for early bilinguals (Sebastián-Gallés & Soto-Faraco, 1999; Pallier, Bosch, & Sebastián-Gallés, 1997). If visual language information is similarly influenced by the age of acquisition effects commonly observed in auditory speech perception tasks, it might be expected that an adult's sensitivity to visual language information will vary depending on when an adult was first exposed to the visual information of a language.

Visual language discrimination tasks can be used to test participants at any age, and provide the opportunity to determine whether sensitivity to visual speech information declines if exposure to a language does not occur until later in life. A reduced sensitivity to visual speech information from a foreign language may have an impact on the ability to learn a second language. To discover whether optimal periods exist for visual language discrimination, it is first imperative to determine when and how the ability to visually discriminate languages develops, and then test how the ability changes as a function of language experience. Comparisons between adults who have learned a language (e.g. English) from birth vs. adults who have learned English as a second language as children or adolescents will yield insight into the possibility of an optimal period for visual language information.

1.5 Infant Visual Perception

Discovering when the ability to discriminate languages using solely visual speech information develops necessitates a few considerations. Although infants can auditorily discriminate and prefer their native language from an unfamiliar language at birth (Mehler et al., 1988; Nazzi et al., 1998), the visual system is not as mature at birth. Furthermore, infants receive and respond to auditory language information in-utero (e.g. Lecanuet et al., 1987; Zimmer et al., 1993), but an infant's first exposure to visual language information does not occur until birth, and even then is at quite a degraded level. Moreover, to process silent talking faces, infants may need to be capable of perceiving detailed facial features, voluntarily converging their eyes to focus on certain features (e.g. the mouth), tracking the features of the face as it speaks, controlling their eyes to scan the face, and attending to the face when interested, but also disengaging attention when bored with the stimuli.

The visual system develops rapidly during the first 6 months of life. At birth, infants' visual acuity is 40 times worse than adults, but between birth and 6 months there is a fivefold increase in acuity followed by slow improvement until 6 years of age (Maurer & Lewis, 2001). Stereoacuity, using retinal disparity to perceive depth, has reliably been observed in 16-26 week old infants and does not improve rapidly until after 12 months of age (Takai, Sato, Tan, & Hirai, 2005). Thus, infants are capable of reliably converging their eyes to focus on a visual display as young as 4 months of age, and their acuity for processing the features of the face develops rapidly by 6 months of age. Since talking faces are not static, the infants must be capable of following an image through motion. Infants' ability to detect directional motion appears at about 7 weeks, but adult levels of motion coherence are not reached until 8-10 years of age (Braddick, Atkinson, & Wattam-Bell, 2003). Infants also need to be able to shift their eyes to examine

different features of the talking face. By 3-4 months of age the cortical eye fields are actively involved in the prospective control of saccades and visual attention (Canfield & Kirkham, 2001). Furthermore, spatial orienting appears to be fairly well established by 6 months of age, with reflexive saccades and inhibition of return showing sizable changes from 2-6 months of age. These facts suggest that infants are capable of controlling their gaze to voluntarily search and focus on visual features of a moving face at 4-6 months of age. Additionally, disengagement of attention improves considerably between 2-4 months of age (Columbo, 2001), which means that infants should be capable of voluntarily switching between two talking faces and able to look away from a talking face during a habituation looking time experiment in order to indicate boredom. Together, these developmental facts suggest that infants are capable of controlling their gaze in order to process the details of moving facial information by 4-6 months of age. However, if this level of detail is not required for visual language discrimination, it may be possible to obtain an effect at an earlier age.

1.6 Languages Differ Visually

It is entirely plausible that infants are sensitive to the visual differences available in talking faces. Languages differ in many ways, and previous research has shown that infants are sensitive to both auditory phonetic and auditory rhythmical language differences. The high correlation between speech sounds and face movements (e.g. Yehia et al., 1998) suggests that both phonetic and rhythmical speech information may be visually represented. Languages can be differentiated according to their phonetic repertoire because phonetic inventories may differ and certain sounds may not be included in some languages. The visual (facial configuration) information may carry much of this segmental phonetic information, representing the vowels and

consonants characteristic of a language (e.g. Benoit, Guiard-Marigny, Le Goff, & Adjoudani, 1996). Languages can also be differentiated according to their rhythmical classes. English, for instance, is commonly referred to as a stress-timed language, whereas languages such as French are syllable-timed (Abercrombie, 1967; Pike, 1945). The visible head, face and jaw movements during speech have been shown to convey rhythmical language information (Munhall & Vatikiotis-Bateson, 1998). Thus, in this thesis, two main language components will be examined, one containing rhythmical information that represents the prosodic patterns of amplitude and duration for each sentence, and the other containing the facial or segmental visual phonetic information, which reflects the facial configurations necessary for producing the sounds characteristic of the language.

1.6.1 Segmental Visual Phonetic (Visegmetic) Cues

Traditionally, the configuration of the lips and mouth that corresponds to the sounds of a language is treated as a visual representation of phonemes, or visemes. However, these only represent a static view of the face producing a language sound. Because the dynamic information in the rest of the face (Vatikiotis-Bateson et al., 1998; Thomas & Jordan, 2004) and even head motions (e.g. Munhall et al., 2004) contribute to the intelligibility of speech, I will use the term "visegmetic" to refer to the dynamically changing arrangement of features on the face (mouth, chin, cheeks, head etc.), basically anything visible on the face and head that accompanies the production of segmental phonetic information. "Visegmetic" segments are therefore meant to capture all facets of motion that the face and head make as language sounds are produced. Thus, visegmetic segments include the time-varying qualities that are roughly equivalent to auditory phonetic segments.

The languages used in this thesis, English and French, share a number of consonants, but there are some visible differences in the phonetic segments characteristic of each language. Although both languages share the consonants [t/d/n/l/s/z] the place of articulation for these consonants can vary between French and English. For example, [t/d/n] are often pronounced with the tongue in a more frontal coronal (dental) position in French, and a more central coronal (alveolar) position in English (Dart, 1998). Furthermore, French contains the consonants [ɲ] and [x], but not the consonants [ŋ/ θ/ʃ/ð̞/ʒ/dʒ]. Of these consonants, the English [ð] may be particularly salient in visual speech as it often requires the placement of the tongue between the teeth, creating a very visible difference between the languages. French also has different vowels, some of which are produced with more lip-rounding which may also be accompanied by differing amounts of protrusion (Benoit & Le Goff, 1998). The aforementioned visible differences between French and English may therefore permit visual language discrimination.

Investigating visegmetic information in infancy is of interest because infants may be capable of tracking the occurrence of visegmetic segments that are characteristic of each language. Infants are capable of using statistical regularities to extract information from the speech signal. For instance, Saffran, Aslin, and Newport, (1996) showed that infants can track transitional probabilities between words, and Maye, Werker and Gerken (2002) demonstrated that infants are sensitive to the statistical frequency of phonetic units in speech streams. Perhaps infants are similarly capable of extracting and tracking the facial configurations unique to, or characteristic of, each language.

1.6.2 Rhythmical Cues

I will use the term "rhythmical" to refer to the perceived rhythm created by the motion of the head, face, mouth and chin during speech. This rhythm may be at the level of the sentence, word, or syllable. I will refer to rhythmical differences as the differences in the patterns of spatial-temporal events in the movements of the head and face, including but not limited to jaw movements, lip opening and closing, and cheek movements.

Most of the world's languages can be roughly classified as stress-timed, syllable-timed (Abercrombie, 1967; Pike, 1945) or mora-timed (Ladefoged, 1975). These classifications are made according to the amplitude, stress and duration differences of the syllables within each word or sentence. The languages used in this thesis are English, which is classified as a stress-timed language and French, which is generally classified as a syllable-timed language (Abercrombie, 1967; Pike, 1945). Stress-timed languages (e.g. English) have some syllables that are generally more intense with a longer duration, while others may be shortened and/or have reduced vowels, creating the characteristic stressed tempo (Bolinger, 1965). The intervals between the stress syllables are usually approximately equal, which means that the syllables in the inter-stress intervals vary according to the number and type of syllable. Syllable-timed languages (e.g. French), however, generally have roughly equal stress placed on each syllable, and each syllable is approximately equal in duration. For French specifically, the final syllable is often lengthened on polysyllabic words and at the ends of sentences or phrases (Wenk & Wioland, 1982), and the first syllable in the utterance may also be stressed (Vaissier, 1983). These features highlight the prosody of the language at the level of the utterance, rather than the word. While English and French may not fit perfectly into stress-timed or syllable-timed

language classifications, they are nonetheless distinctive in their time-varying patterns of spatial-temporal events, and rhythmically distinct for the purposes of the studies in this thesis.

To better understand the rhythmical differences between languages, Ramus, Nespor and Mehler (1999) calculated how language rhythm may be classified using the duration of vocalic and consonantal intervals. Given that the movements of the mouth, face and head can be used to recover a significant portion of the sound signal (e.g. Munhall & Vatikiotis-Bateson, 1998; Yehia et al., 1998), the percentage of consonant and vowel durations may also be visually represented. Ramus and colleagues' calculations for language rhythm were derived using three variables: proportion of vocalic intervals, standard deviation of the duration of vocalic intervals within each sentence, and the standard deviation of the duration of consonant intervals within each sentence. The calculations of Ramus and colleagues show how language rhythm can be influenced by contrastive vowel length (e.g. Japanese, Dutch), vowel lengthening in specific contexts (e.g. Italian), long vowels (e.g. English, Dutch, French), and the number of consonants. As a language gains more consonant clusters, it also tends to gain more stressed syllables and more vowel reduction (e.g. English, Dutch, Catalan). If the proportion of vowel and consonant information regulates language rhythm, features in the face producing these language sounds may therefore convey the rhythmical language information. Munhall and Vatikiotis-Bateson (1998) have shown that there is a strong correlation between the dynamic auditory signal amplitude and the corresponding motion expressed by multiple sites on the face and head. The visual language information conveyed by a speaker's head should therefore capture and represent rhythmical language information, and these visual rhythmical differences may then be used to discriminate languages visually.

French and English fall into two separate rhythmical categories, increasing the likelihood that they will be discriminable visually. However, the findings from this thesis, with regard to rhythmically distinct languages, cannot necessarily be generalized to other languages from the same rhythmical categories. As there is a high degree of variability within rhythmical categories, the differences between French and English may not hold for other stress-timed and syllable-timed languages.

1.6.3 Visegmetic versus Rhythmical Cues

To better understand the cues that may be used for visual language discrimination, manipulations of the speech signal were designed to separate the rhythmical from the visegmetic cues in order to determine the conditions under which these cues are sufficient for language discrimination.

To assess the possibility that infants can track differences in facial configurations, the visual speech signal was manipulated to isolate visegmetic information, and rhythmical language cues were removed. Studies using the auditory speech stream disrupt the rhythmical speech cues in sentences by randomly reordering the words of the sentence (Dehaene-Lambertz & Houston, 1998). To accomplish this visually in my experiments, the visual speech stream was cut into 200 msec segments. The speech stream was cut below the level of the average syllable length for each language. French generally has shorter syllable lengths, so it was important to cut at a point below the length of the average syllable for both languages. The average syllable length for each sentence was calculated by removing the silence from the auditory speech stream and dividing by the number of syllables for each sentence. Because the average syllable length of the French sentences used for the experiments in this thesis is 223 msec and the average syllable length for

the English sentences is 236 msec, the sentences were cut into 200 msec segments in order to prevent rhythmicity at the syllable level from influencing the infants' language discrimination. The segments were then randomly reordered for each sentence in each language. The original sentences were filmed with the speakers' heads held quite still, so although the transitions between the random 200 msec clips were somewhat choppy, they blended surprisingly well. The randomly reordered segments caused a rhythmical disruption for all of the sentences, regardless of the language, while maintaining some of the forward or natural motion of the visegmetic face information.

The clips were also reversed so that both rhythmical language and segmental visual phonetic cues were distorted. Although reversing the sentences disrupts the rhythmicity of the speech signal, it is a consistent disruption across each language, so infants may none-the-less be able to discriminate the languages using solely reversed language rhythmical information.

Determining the visual speech cues that facilitate language discrimination is also important for understanding what visual properties of the speech signal are influenced by optimal periods for visual language experience. Understanding which visual speech properties become less discriminable by infants will provide insight into the visual speech cues that second language learning adults have difficulty perceiving. This in turn will help explain the challenges adults encounter when learning a second language.

1.7 Visual Preference for the Native Language

Auditorily, newborns have been found to prefer listening to passages recited in their native language (Moon, Cooper, & Fifer, 1993). We wanted to examine whether infants also prefer watching sentences recited in their native language. Infants' abilities to process moving

facial stimuli by 4 months, and the rapid increase in visual acuity during the first 6 months, makes it likely that infants are able to visually process and remember the movements and changes in facial configuration that accompany visual speech by 4-6 months. Although infants appear to process facial details sufficiently by 4 months, many visual abilities such as visual acuity stabilize around 6 months, so the first study tested infants at 6 months. Newborns only auditorily discriminate between rhythmically dissimilar languages. French and English were therefore chosen because they differ on both rhythmical and phonetic properties.

Infants' visual preference for their native language was tested at 6 months. Fantz (1956) demonstrated the benefit of using infants' looking times to assess their preference for images that are placed side-by-side. We chose not to use side-by-side images because MacKain, Studdert-Kennedy, Spieker, and Stern (1983) tested 6-month-olds with a language study, and found a strong right-side bias for disyllables. Although they interpreted this as evidence for the left lateralization of language, the side bias could mask a preference for one language over the other. Sequential, central presentation, preference designs with visual talking faces have successfully been used with infants as young as 7 weeks (Cooper & Aslin, 1990; Pegg, Werker, & McLeod, 1992). Thus, a sequential preference procedure was used for this experiment.

Twenty English monolingual infants were shown *silent* video clips of bilingual (French/English) speakers reciting sentences in both languages. The sentences were from a subset of adult-directed silent video clips of speakers reciting sentences in English and French from the children's story *The Little Prince*. Sentences from the same collection of video clips were used for all of the experiments in this thesis and are listed in Appendix A.

Infants were tested while sitting on their parent's lap facing a television screen in a sound attenuated room. The parents wore blackened sunglasses to prevent them from viewing the clips

and potentially influencing their infant. The clips were presented on the television screen individually at mid-line, and alternated between French and English sentences. The infants' looking times to the English vs. the French sentences were measured. Overall, there was no distinct preference for the native language, but there was an effect for the first clip that the infants watched. If the very first clip was the familiar language English, the infants watched the clip significantly longer than infants who watched the first clip in the unfamiliar language French (Weikum, Werker, Vouloumanos, Soto-Faraco, & Sebastián-Gallés, 2005). Although this was not strong evidence for a preference, the initial bias for watching the familiar native language longer was suggestive of an ability to visually discriminate the languages.

1.8 Hypotheses

       The previous research described suggests that both infants and adults may be capable of discriminating languages visually. This thesis therefore tests the following hypotheses:

Sensitivity to visual language information:

-Infants and adults are sensitive to visual language information and can use this information to discriminate their native language from an unfamiliar language.

Experience related changes in sensitivity to visual language information:

Experience with faces and auditory speech sounds results in perceptual (re)organization during the first year of life.

-Infants' abilities to discriminate languages change to reflect their experience with the languages tested.

-Adults' sensitivity to visual language information reflects their early experience (language known or unknown and age learned) with the language(s) tested.

Changes in sensitivity to specific visual speech cues affects performance at different ages:

-Infants are sensitive to the segmental visual phonetic "visegmetic" and rhythmical cues that characterize rhythmically distinct languages.

-Infants are sensitive to different visual speech cues (visegmetic or rhythmical) at different ages depending on their language-learning environment.

1.9 Thesis Chapter Outline

  The following chapters in the thesis address these hypotheses and continue to examine the role of visual speech information in language discrimination. Chapter II examines whether infants can discriminate their native language from an unfamiliar language, and how this changes with language experience over the first year of life. Chapter III examines whether changes in sensitivity to visual language information in infancy has an impact on adult visual language discrimination. Adults with different early language experience were tested to determine if adult visual language discrimination proficiency is influenced by the age at which a language was learned in early life. The experiments in Chapter IV test the impact of two different manipulations of the visual speech clips. Rhythmical and segmental cues were isolated or distorted in order to determine the cues that are necessary for, or facilitate language discrimination in infancy. Success or failure on language discrimination tasks with these isolated speech cues suggests the components of the visual speech signal that may be affected by optimal periods for visual language information. Finally, Chapter V discusses how these studies relate to each other, their limitations, and future directions.

  The findings discussed in this thesis further our understanding of how infants come to recognize their native language(s) visually, and also provide a platform for further investigation. Future experiments will examine not only the initial capabilities of the infants to process multimodal speech information, but how the brain continually organizes to accommodate the auditory and visual speech information characteristic of an infant's language experience.

1.10 References

Abercrombie, D. (1967). *Elements of General Phonetics*. Chicago, IL: Aldine Publishing

   Company.

Baker, S. A., Golinkoff, R. M., & Petitto L. (2006). New insights into old puzzles from infants'

   categorical discrimination of soundless phonetic units. *Language Learning and*

   Development 2(3), 147-162.

Benoit, C., Guiard-Marigny, T., Le Goff, B., & Adjoudani, A. (1996). Which components of the

   face do humans and machines best speechread? In D. Stork & M. E. Hennecke (eds.),

   Speechreading by Man and Machine (pp. 315-328). Germany: Springer-Verlag.

Benoit, C., & Le Goff, B. (1998). Audio-visual speech synthesis from French text: eight years of

   models, designs and evaluation at the ICP. *Speech Communication, 26*, 117-129.

Bernstein, L. E., Auer, E. T., Moore, J. K., Ponton, C. W., Don, M., & Singh, M. (2002). Visual

   speech perception without primary auditory cortex activation. *Neuroreport, 13*(3), 311-

   315.

Bernstein, L. E., Demorest, M. E., & Tucker, P. E. (2000). Speech perception without hearing.

   Perception & Psychophysics, *62*, 233-252.

Best, C. T., McRoberts, G. W., LaFleur, R., & Silver-Isenstadt, J. (1995). Divergent

   developmental patterns for infants' perception of two nonnative consonant contrasts.

   Infant Behavior and Development, 18(3), 339-350.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual

   reorganization for nonnative speech contrasts: Zulu click discrimination by English-

   speaking adults and infants. *Journal of Experimental Psychology: Human Perception and*

   Performance, 14, 345-360.

Bolinger, D. (1965). Pitch accent and sentence rhythm. In I. Abe & T. Kanekiyo (eds.) *Forms of English: Accent, Morpheme, Order* (pp.139-180). Cambridge, MA: Harvard University press.

Braddick, O., Atkinson, J., & Wattam-Bell, J. (2003). Normal and anomalous development of visual motion processing: motion coherence and 'dorsal stream vulnerability'. *Neuropsychologia, 41*(13), 1769-1784.

Brutt-Griffler, J. & Varghese, M. (2004). Introduction. *International Journal of Bilingual Education and Bilingualism, 7*(2-3), 93-101.

Burnham, D., & Dodd, B. (2004). Auditory-visual speech integration by prelinguistic infants: Perception of an emergent consonant in the McGurk effect. *Developmental Psychobiology, 45*, 204-220.

Burns, T. C., Yoshida, K. A., Hill, K., & Werker, J. F. (2007). The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics, 28*(3), 455-474.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire P. K., Woodruff, P. W., Iversen, S. D., & David, A. S. (1997). Activation of auditory cortex during silent lipreading. *Science, 276*, 593-596.

Canfield, R. L., & Kirkham, N. Z. (2001). Infant cortical development and the prospective control of saccadic eye movements. *Infancy, 2*(2), 197-211.

Columbo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology, 52*, 337-367.

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech within the first month after birth. *Child Development, 61*, 1584-1595.

Dart, S. (1998). Comparing French and English coronal consonant articulation. *Journal of Phonetics, 26,* 71-94.

Dehaene-Lambertz, G., & Houston, D. (1998). Language discrimination response latencies in two-month-old infants. *Language and Speech, 41*, 21-43.

Desjardins, R. N., & Werker, J. F. (2004). Is the integration of heard and seen speech mandatory for infants? *Developmental Psychobiology*, *45*, 187-203.

Dodd, B. (1979). Lip-reading in infants: Attention to speech presented in- and out-of-synchrony. *Cognitive Psychology, 11*, 478-484.

Dodd, B., & Burnham, D. K. (1988). Processing speechread information. *The Volta Review: New Reflections on Speechreading*, *90*, 45-60.

Fantz, R. L. (1956). A method for studying early visual development. *Perceptual and Motor Skills, 6*, 13-15.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *The Journal of the Acoustical Society of America, 89*, 395-411.

Gagne, J. P., Masterson, V., & Munhall, K. G. (1994). Across talker variability in auditory, visual, and audiovisual speech intelligibility for conversational and clear speech. *Journal of the Academy of Rehabilitative Audiology, 27*, 135-158.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston, MA: Houghton Mifflin.

Green, K. P. (1998). The use of auditory and visual information during phonetic processing:

Implications for theories of speech perception. In R. Campbell & B. Dodd (eds.).

Hearing by Eye II: Advances in the Psychology of Speechreading and Audiovisual

Speech (pp.3-25). Hove, UK: Psychology Press.

Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood.

Psychological Science, 16, 48-55.

Helfter, K. S. (1997). Auditory and audio-visual perception of clear and conversational speech.

Journal of Speech, Language and Hearing Research, 40, 432-443.

Hensch T. K. (2004). Critical period regulation. *Annual Review Neuroscience. 29*, 549-579.

Hollich, G., Newman, R. S., & Jusczyk, P. W. (2005). Infants' use of synchronized visual

information to separate streams of speech. *Child Development, 76*(3), 598-613.

Kamachi, M., Hill, H., Lander, K., & Vatikiotis-Bateson, E. (2003). Putting the face to the voice:

matching across modality. *Current Biology, 13*(19)**,** 1709–1714.

Kuhl, P. K., & Meltzoff, A. N. (1984). The intermodal representation of speech in infants.

Infants Behavior and Development, 7, 361-381.

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), F1–F9.

Kuhl, P. K., Tsao F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of

short-term exposure and social interaction on phonetic learning. *Proceedings of the*

National Academy of Sciences, *100*, 9096-9101.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic
experience alters phonetic perception in infants by 6 months of age. *Science, 255,* 606-
608.

Ladefoged, P. (1975). A Course in Phonetics. New York: Harcourt Brace Jovanovich.

Lecanuet, J. P., Granier-Deferre, C., DeCasper, A. J., Maugeais, R., Andrieu, A. J., & Busnel, M.
C. (1987). Perception and discrimination of language stimuli. Demonstration from the
cardiac responsiveness. Preliminary results. *Proceedings of the National Academy of
Sciences*, *Paris, III*, 161-164.

Lenneberg, E. H. (1967). *The biological foundations of language*. New York, NY: Wiley.

Lewkowicz, D. J., & Ghazanfar, A. A. (2006). The decline of cross-species intersensory
perception in human infants. *Proceedings of the National Academy of Sciences of the
USA, 103,* 6771-6774.

MacKain, K., Studdert-Kennedy, M., Spieker, S., & Stern, D. (1983). Infant intermodal speech
perception is a left-hemisphere function. *Science, 219*, 1347-1348.

MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production.
Behavioral and Brain Sciences, 21, 499-546.

MacSweeney, M., Amaro, E., Calvert, G. A., Campbell, R., David, A. S., McGuire, P., Williams,
S. C., Woll, B., & Brammer, M. J. (2000). Silent speechreading in the absence of scanner
noise: an event-related fMRI study. *Neuroreport. 11*(8), 1729-1733.

Massaro, D. W. (1987). *Speech perception by ear and by eye: A paradigm for psychological
inquiry*. Hillsdale, NJ: Erlbaum.

Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for
perceptual reorganization. *Infancy, 10*(3), 241-265.

Maurer, D., & Lewis, T. L. (2001). Visual Acuity: The role of visual input in inducing postnatal change. *Clinical Neuroscience Research, 1*(4), 239-247.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*, B101-B111.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 46-748.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29*, 143-178.

Mills, A. E. (1983). *Language Acquisition and the Blind Child.* London: Croom Helm.

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

Munhall, K. G., Jones, J. A., Callan, D. E., Kuratate, T., & Vatikiotis-Bateson, E. (2004). Visual prosody and speech intelligibility - Head movement improves auditory speech perception. *Psychological Science, 15*(2), 133-137.

Munhall, K. G., & Vatikiotis-Bateson, E. (1998). The moving face during speech communication. In R. Campbell, B. Dodd & D. Burnham (eds.) *Hearing by Eye II: Advances in the Psychology of Speechreading and Auditory-visual Speech* (pp. 123-139). East Sussex, UK: Psychology Press Ltd.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception & Performance, 24*(3), 756-766.

Olson, I. R., Gatenby, J. C., & Gore, J. C. (2002). A comparison of bound and unbound audio-visual information processing in the human cerebral cortex. *Brain Research Cognitive Brain Research 14*, 129-138.

Pallier, C., Bosch, L., & Sebastián-Gallés, N. (1997). A limit on behavioral plasticity in speech

    perception. *Cognition, 64,* B9–B17.

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the

    first year of life? *Science*, *296*, 1321-1323.

Pascalis, O., Scott, L. S., Kelly, D. J., Shannon, R. W., Nicholson, E., Coleman, M., & Nelson,

    C. A. (2005). Plasticity of face processing in infancy. *Proceedings of the National*

    Academy of Science of the USA, *102,* 5297-5300.

Patterson, M. L., & Werker, J. F. (2003). Two-month-old infants match phonetic information in

    lips and voice. *Developmental Science, 6*, 191-196.

Pegg, J. E., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-

    directed speech: evidence from 7-week-old infants. *Infant Behavior and Development,*

    15(3), 325-345.

Pike, K. (1945). *The Intonation of American English*. Ann Arbor, MI: University of Michigan

    Press.

Piske, T., MacKay, I.R.A., Flege, J.E. (2001). Factors affecting degree of foreign accent in an

    L2: a review. *Journal of Phonetics, 29*, 191-215.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of non-native vowel

    contrasts. *Journal of Experimental Psychology: Human Perception and Performance*,

    20(2), 421-435.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal.

    Cognition, 73(3), 265-292.

Reisberg, D., McLean, J., & Goldfield, A. (1987). Easy to hear but hard to understand: A lip-reading advantage with intact auditory stimuli. In B. Dodd and R. Campbell (Eds.), *Hearing by eye: The psychology of lip reading* (pp. 97-114). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rosenblum, L. D. (2005). The primacy of multimodal speech perception. In D. Pisoni & R. Remez (Eds.), *Handbook of speech perception* (pp. 51–78). Malden, MA: Blackwell.

Rosenblum, L. D., Miller, R. M., & Sanchez, K. (2007). Lip-read me now, hear me better later: cross-modal transfer of talker-familiarity effects. *Psychological Science 18*(5), 392–396.

Rosenblum, L. D., Schmuckler, M. A., & Johnson, J. A. (1997). The McGurk effect in infants. *Perception & Psychophysics, 59*(3), 347-357.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926-1928.

Sebastián-Gallés, N., & Soto-Faraco, S. (1999). Online processing of native and nonnative phonemic contrasts in early bilinguals. *Cognition, 72*, 111–123.

Soto-Faraco, S., Navarra, J., Weikum, W. M., Vouloumanos, A., Sebastián-Gallés, N., & Werker, J. F. (2007). Discriminating languages by speechreading. *Perception and Psychophysics, 69,* 218-231.

Streeter, L. A. (1976). Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature, 259,* 39-41.

Sullivan, J. W., & Horowitz, F. D. (1983). The effects of intonation on infant attention: the role of the rising intonation contour. *Journal of Child Language, 10*(3), 521-534.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America, 26*(2), 212-215.

Summerfield, A. Q. (1979). Use of visual information in phonetic perception. *Phonetica, 36*, 314-331.

Summerfield, A. Q. (1987). Some preliminaries to a comprehensive account of audiovisual speech perception. In B. Dodd and R. Campbell (Eds.), *Hearing by eye: The psychology of lipreading* (pp. 3-51). Hillsdale, NJ: Erlbaum.

Takai, Y., Sato, M., Tan, R., & Hirai, T. (2005). Development of Stereoscopic Acuity: Longitudinal study using a computer-based random-dot stereo test. *Japanese Journal of Opthamology, 49*(1), 1-5.

Tees, R. C. (2001). Critical and sensitive periods. In P. Winn (Ed.), *Dictionary of biological psychology* (pp. 195, 701). London: Routledge Press.

Thomas S. M., & Jordan T. R. (2004). Contributions of oral and extraoral facial movement to visual and audiovisual speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 873-888.

Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development, 47*, 466-472.

Vaissier, J. (1983). Language-independent prosodic features. In A. Cutler and D. R. Ladd (Eds.), *Prosody: Models and Measurements* (pp. 53-66). Berlin: Springer-Verlag.

Vatikiotis-Bateson, E., Eigsti, I. M., Yano, S., & Munhall, K. G. (1998). Eye movement of perceivers during audiovisual speech perception. *Perception & Psychophysics, 60,* 926-940.

Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, *10*(2), 159-164.

Vroomen, J., & de Gelder, B. (2000). Sound enhances visual perception: cross-modal effects of

auditory organization on vision. *Journal of Experimental Psychology: Human Perception

and Performance 26*(5), 1583-1590.

Weikum, W. M., Werker, J. F., Vouloumanos, A., Soto-Faraco, S., & Sebastián-Gallés, N.

(2005, April). Silent talking heads cue language discrimination in infancy. *Society for

Research in Child Development (SRCD)*, Atlanta, Georgia.

Wenk, B., & Wioland, F. (1982) Is French really syllable-timed? *Journal of Phonetics, 10*, 193-

216.

Werker, J. F., Gilbert, J. H. V., Humphrey, G. K., & Tees, R. C. (1981). Developmental aspects

of cross-language speech perception. *Child Development*, *52*, 349-355.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual

reorganization during the first year of life. *Infant Behavior and Development*, *7*, 49-63.

Werker J. F., & Tees R. C. (1999). Influences on infant speech processing: Toward a new

synthesis. *Annual Review of Psychology, 50*, 509-535.

Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity

and commitment in language systems of the brain. *Developmental Psychobiology*, *46*(3),

233-251.

Yehia, H. C., Rubin, P. E., & Vatikiotis-Bateson, E. (1998). Quantitative association of vocal-

tract and facial behavior. *Speech Communication, 26*, 23-43.

Zimmer, E. Z., Fifer, W. P., Kim, Y. I., Rey, H. R., Chao, C. R., & Meyers, M. M. (1993).

Response of the premature fetus to stimulation by speech sounds. *Early Human

Development, 33*, 207-215.

CHAPTER II – Visual Language Discrimination in Infancy[2]

2.1 Introduction

Talking faces are among the most dynamic and salient stimuli available to infants, and the facial movements accompanying speech influence adult (Sumby & Pollack, 1954) and infant (Dodd & Burnham, 1988) speech perception. Recently it was reported that facial speech information alone is sufficient for language discrimination in adults (Soto-Faraco et al., 2007). Though it is well established that young infants can discriminate languages auditorily (Mehler et al., 1988; Bosch & Sebastián-Gallés, 1997), it is unknown whether infants can discriminate languages visually. We examined whether 4-month-olds can visually distinguish their native language (English) from an unfamiliar language (French). Thus, because exposure to specific auditory and visual information in infancy is essential for maintaining many early appearing native language (Werker & Tees, 1984; Kuhl et al., 2006; Lewkowicz & Ghazanfar, 2006), musical (Hannon & Trehub, 2005), and face perception sensitivities (Pascalis et al., 2005) we compared monolingual English infants to French/English bilingual infants at 6 and 8 months.

2.2 Method

Discrimination was tested using silent video clips of three bilingual French/English speakers reciting sentences in each language. Every trial contained a video clip of a different sentence, by one speaker in one language. The infants ($n = 36$) were presented with video clips from one of the languages until their looking time declined to a 60% habituation criterion. Test trials using the same speakers but different sentences from the other language were shown to examine whether the infants' looking time increased, indicating that they noticed the language

---

[2] A version of this chapter has been published. Weikum, W. M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., & Werker, J. F. (2007). Visual Language Discrimination in Infancy. *Science, 316,* 1159. http://www.sciencemag.org/cgi/content/abstract/316/5828/1159

change. The test trials where the language was switched were compared to a control condition (*n* = 36) for which the test trials were always different sentences but in the same language as the habituation trials.

2.3 Results

A repeated-measures analysis of variance (ANOVA) including age (4, 6 or 8 months), condition (language switch vs. control), and trial (habituation vs. test) revealed only a significant three-way interaction [$F(2, 66) = 3.71, p < .05$] Simple main effects analyses showed that the infants looked significantly longer at the language switch test trials (Fig. 2.1), compared to the control trials, at 4 months [$F(1, 22) = 4.70, p < .05$] and 6 months [$F(1, 22) = 4.19, p = .05$], but not at 8 months [$F(1, 22) = 1.18, p = .29$].

Monolingual Infants at 4, 6 and 8 months

Figure 2.1. Mean looking time in seconds to silent talking faces. The y axis represents infant looking time; the x axis represents the trial that the infant was shown (final habituation trials or test trials). Error bars represent the standard error of the mean. Experimental (language switch) and control (language same) conditions for monolingual infants at 4, 6 and 8 months.

The finding that infants can visually discriminate their native language from an unfamiliar language at 4 and 6 months, but not at 8 months parallels declines in performance seen in other perceptual domains. Indeed, across the first year of life, infants' performance declines on the discrimination of non-native consonant and vowel contrasts (Werker & Tees, 1984, Kuhl et al., 2006), non-native musical rhythms (Hannon & Trehub, 2005), cross-species individual faces (Pascalis et al., 2005) and cross-species face/voice matching (Lewkowicz & Ghazanfar, 2006). Thus, it appears that specific experience is necessary for maintaining sensitivity to some initial perceptual discriminations. To determine if regular exposure to both French and English confers an advantage in visual language discrimination, we compared bilingual French/English infants (n = 24) to the monolingual English counterparts that we had tested previously at 6 and 8 months.

At 6 months, a 2 x 2 repeated-measures ANOVA analyzing language group (monolingual vs. bilingual) and trial (habituation vs. test) yielded a significant effect for trial [$F(1, 22) = 6.652$, $p < .02$] with no interaction. A similar analysis at 8 months yielded only a significant trial by condition interaction [$F(1, 22) = 6.92$, $p < .02$]. Simple main effects analyses of this interaction showed that at 8 months, only the bilingual infants looked significantly longer to the change in language [$F(1, 11) = 7.1$, $p < .05$; Fig. 2.2].

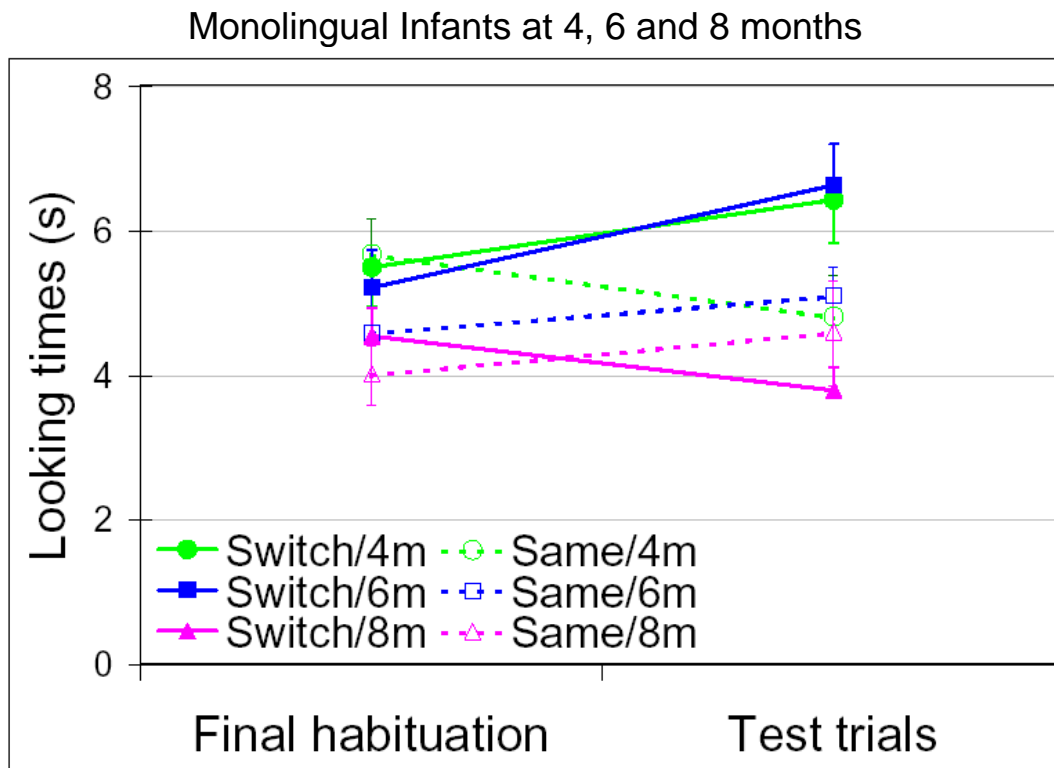Monolingual and Bilingual Infants at 6 and 8 months
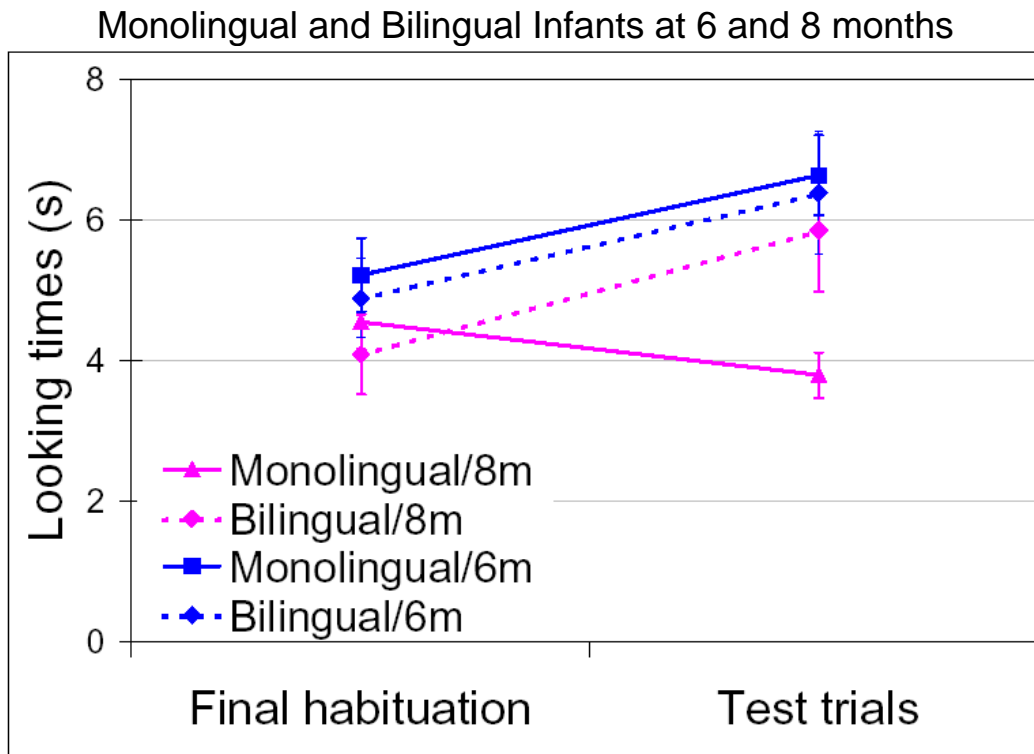
Figure 2.2. Mean looking time in seconds to silent talking faces. The y axis represents infant looking time; the x axis represents the trial that the infant was shown (final habituation trials or test trials). Error bars represent the standard error of the mean. Experimental conditions for monolingual (re-plotted from Fig. 2.1) and bilingual infants at 6 and 8 months.

2.4 Discussion

Traditionally, visual speech has been regarded as a redundant signal in verbal communication, thereby suggesting that its influence is to merely facilitate speech perception in the presence of an auditory signal. This study shows that visual speech information alone (without the aid of auditory information) is sufficient for language discrimination in infancy. Visual speech may therefore be important for helping infants differentiate between languages as

they learn them from a very early age (as young as 4 months according to our findings).

Moreover, these findings indicate that visual speech may also play a more critical role than

previously anticipated in helping infants narrow their perceptual sensitivities. In addition to a

decline in the ability to perceive non-native auditory speech and face identity information,

infants also experience a decline in the ability to discriminate languages visually (unless they are

familiar with both languages). This emphasizes how visual speech sensitivities evolve to support

the perception of visual distinctions necessary in an infant's language-learning environment.

Notably, bilingual infants advantageously maintain the discrimination abilities needed for

separating and learning multiple languages. Thus, before infants recognize most of the words

characteristic of their native language, they may be capable of using solely visual speech cues to

recognize and discriminate between speakers of their native language(s).

2.5 Materials & Methods

2.5.1 Participants

The final sample included 96 healthy, full-term infants. Additional infants were excluded

for failure to habituate (7), fussiness (15), parental interference (7), distraction

(3), and experimenter error (1). Each condition, [monolingual, bilingual (exposed to at least 25%

English and French according to parental estimates) and control] included 12 infants per age.

Infants were tested at 4 months (3.87–5.17m, M=4.5m), 6 months (5.97–7.17m, M=6.57m) and 8

months (7.77–9.2m, M=8.43m).

2.5.2 Stimuli

Silent video clips were recorded using three bilingual French/English speakers who recited sentences from *The Little Prince/Le Petit Prince* (for sentence details see Appendix A). Sentences from both languages were recited by each speaker. Each clip consisted of one bilingual speaker reciting a different sentence in either English or French. A colourful expanding and retracting ball attracted the infant's attention toward the screen for the start of every trial.

2.5.3 Procedure

The infants were seated on their parent's lap in a sound attenuated room approximately 4 feet from a 27-inch TV screen. The parents wore blackened sunglasses to prevent them from viewing the visual clips and potentially influencing their infant. The experimenter controlled the study from a separate room, and watched the infant's looking responses via a closed circuit camera. The experimenter pressed a key whenever the infant looked at the stimuli. The presentation of the stimuli, storage of online looking time, and calculation of the habituation criterion were run using Habit 2000 software (Cohen, Atkinson, & Chaput, 2002). The experimenter was thus blind to the timing of the change from habituation to test trials. The infants were habituated by successively presenting clips (each clip contained a different sentence) from one of the languages until the infant's looking time across three trials declined to a preset criterion of 60%. Choosing a criterion maintained that the infants were generally bored or familiarized with the stimuli to a similar degree. Half the infants were then shown the same language (control condition) and half were shown the opposite language (switch condition). Each trial lasted a maximum of 16 seconds, but terminated when the infant

looked away for more than 2 seconds (indicating boredom with the particular clip). The trials

were organized into blocks of three so that the infants always watched the same speakers in the

same order across the habituation and test trials. The infants could habituate in as few as six

trials, but a maximum of 24 trials were available for habituation. The three test trials started

seamlessly after the habituation criterion was reached. The test trials were repeated once to

ensure that the infants had enough exposure to the clips in order to notice the subtle language

switch. All test trial statistics used an average of the six test trials for each infant. The control

infants (half tested with French, half with English) were tested on new sentences from the same

language presented during habituation. The test trials were compared to the final three

habituation trials. If the infants' looking times recovered during the test trials when the language

switched, it was taken as an indication that they noticed the language switch. Videos of each

infant were digitized and coded off-line, frame-by-frame, to obtain precise looking times for the

analyses. This frame-by-frame coding allowed for a check of online coding that may have

resulted in bias; no evidence of this was found.

2.6 References

Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition, 65*, 33-69.

Cohen, L. B., Atkinson, D. J., & Chaput, H. H. (2002). *Habit 2002: A new program for obtaining and organizing data in infant perception and cognition studies* (version 1.0) [Computer Software], Austin, TX: The University of Texas.

Dodd, B., & Burnham, D. K. (1988). Processing speechread information. *The Volta Review: New Reflections on Speechreading*, *90*, 45-60.

Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science, 16*, 48-55.

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), F1–F9.

Lewkowicz, D. J., & Ghazanfar, A. A. (2006). The decline of cross-species intersensory perception in human infants. *Proceedings of the National Academy of Sciences, 103*, 6771-6774.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29*, 143-178.

Pascalis, O., Scott, L. S., Kelly, D. J., Shannon, R. W., Nicholson, E., Coleman, M., & Nelson, C. A. (2005). Plasticity of face processing in infancy. *Proceedings of the National Academy of Science of the USA*, *102,* 5297-5300.

Soto-Faraco, S., Navarra, J., Weikum, W. M., Vouloumanos, A., Sebastián-Gallés, N., &
Werker, J. F. (2007). Discriminating languages by speechreading. *Perception and Psychophysics, 69,* 218-231.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America, 26*(2), 212-215.

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

CHAPTER III- Visual Language Discrimination in Adults[3]

3.1 Introduction

The majority of the world's population is bilingual (Brutt-Griffler & Varghese, 2004), but anyone who has tried to learn a second language, especially as an adult, can attest to the challenges of this task. Difficulties encountered when learning a second language in adulthood may be a consequence of missing input during optimal periods for speech perception that exist in early infancy and childhood (Werker & Tees, 2005). Studies examining optimal periods for speech perception have typically focused on the auditory speech stream as it relates to monolingual speakers. However, visual speech information has been found to dramatically improve (Sumby & Pollack, 1954) and influence (McGurk & MacDonald, 1976) speech intelligibility. Furthermore, second language learners have been found to perceive visual speech information differently than their monolingual counterparts (Hardison, 1996, 2003; Navarra & Soto-Faraco, 2007). This research therefore investigates optimal periods for exposure to visual speech by examining whether information available in a speaker's face is influenced by the age at which a second language is learned.

Studying speech perception using solely visual means is problematic due to the high degree of variability and generally poor lipreading abilities of adults. (e.g. Bernstein, Demorest, & Tucker, 2000). Recently, however, a visual task has been developed that allows researchers to examine whether or not solely visual speech information can be used to discriminate languages. Soto-Faraco et al. (2007) established that monolingual and bilingual adults are capable of visually distinguishing different languages, but they must know one or both languages. The study also showed that it is statistically unlikely that word recognition facilitates the language

---

[3] A version of this chapter will be submitted for publication. Weikum, W. M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., & Werker, J. F. Visual language discrimination in adults.

discrimination of the adults. Infants tested on a variation of this task discriminate their native language from an unfamiliar language as young as 4 months, but this ability declines by 8 months unless the infants are familiar with both languages (Weikum et al., 2007). The effect of language experience on the maintenance of both adult and infant visual language discrimination suggests that infancy is an important period for exposure to visual language information.

Previous studies examining auditory speech perception and production have suggested that age 6 may be an important cut-off for phonological processing and accent-free speech (e.g. Flege & Fletcher, 1992; Flege, Munro, & Mackay, 1995). Studies have also shown that even early bilinguals may show differences on difficult phonological tasks (Pallier, Bosch, & Sebastián-Gallés, 1997; Sebastián-Gallés & Soto-Faraco, 1999). Because an effect has been found for visual language discrimination by 8 months of age (Weikum et al., 2007), we were interested to determine whether this provides evidence for an optimal period in infancy that has life long consequences, or whether it shows a (re)organization process that has begun, but has not yet become permanent. The following research therefore examines whether exposure to visual language information in infancy (before age 2 years) compared to early childhood (age 2-6 years) and late childhood (over 6 years) has an influence on adults' performance on visual language discrimination tasks. We used a broad range (0-2 years) to cover infancy because adults are not accurate in reporting precisely when input from a second language began (especially if it was early in life). Although a cut-off at 6 months of age would have provided an ideal comparison for the perceptual change found in the infant work, to be conservative and accurate we used a 2 year cut-off.

3.2 Method

Adults who had learned English from birth (L1), or had not learned English from birth, but had learned English as a second language (L2) participated. Those who had learned another language (not French) from birth were further divided to examine age effects. English learned with another language in infancy (age 0-2 years), English learned as a second language in early childhood (age 2-6 years), and English learned as a second language in late childhood (age 6+ years). None of the adults were fluent in French, but some had had exposure in school.

The adults watched silent video clips that were created by recording the faces of three bilingual (English/French) speakers reciting sentences in both languages. The silent video clips were presented sequentially in pairs that contained the same speaker, but different sentences in the same or different languages. Following the presentation of both clips, the adults indicated whether the two clips were in the "same" language or "different" languages. The participants watched 24 pairs of the silent video clips that were presented in either a "random" or "blocked" design. In the "random" condition, the participants randomly saw one of the speakers on each trial. In the blocked set, all clips from each speaker were blocked together to allow the participant to develop familiarity with each speaker. This allowed for a test of potential improvement across exposure to each speaker. The order of the speakers was counterbalanced for each condition.

3.3 Results

The adults correctly identified the clips significantly better than chance in both the L1 [M=60%, $t(59) = 6.84$, $p < .001$] and L2 [M=54%, $t(59)=3.00$, $p < .05$] groups and the Random

[*M*=57%, *t*(59) = 4.56, *p* <.001] and Block [*M*=58%, *t*(59) = 4.99, *p* <.001] conditions. A

univariate analysis of variance (ANOVA) analyzing sex, language (L1 or L2), and condition

(block or random) yielded only a significant main effect for language [$F(1,119) = 8.08$, *p* < .05;

Fig 3.1]. Simple main effect analyses showed that the L2 speakers performed significantly more

poorly than the English L1 speakers [$F(1,119) = 5.40$, *p* < .05].

**Accuracy for English learned from Birth (L1)
or after Age 2 Years (L2)**
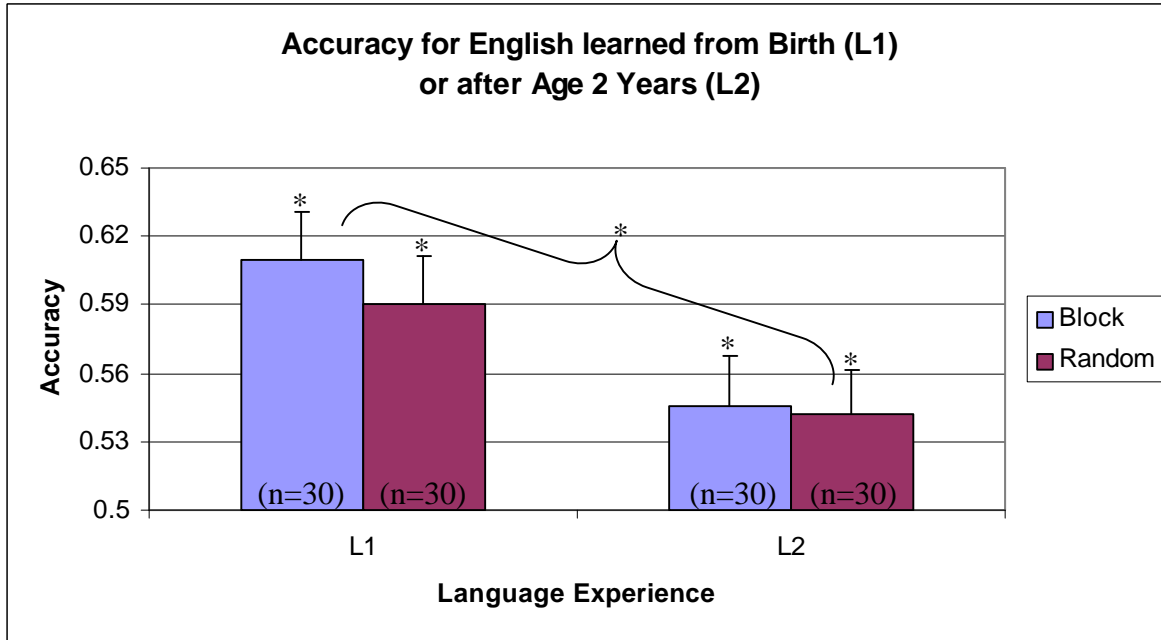
Figure 3.1. Accuracy of adults in both Random and Block condition. The y axis represents

mean accuracy; the x axis represents whether the adults had learned English from birth (L1) or

after the age of 2 years (L2). Error bars represent the standard error of the mean. *p <.05.

To probe for an optimal period in visual language information, all the adults tested were

divided according to those who had acquired English from birth (0-2 years), English in early

childhood (2-6 years), and English in late childhood (after age 6 years). An ANOVA analyzing

age English was learned (birth, age 2-6, over 6), yielded a significant effect for age [$F(2,117) =$

5.55, *p* < .05]. Planned comparisons showed that the multi-lingual adults who had learned

English in infancy (0-2 years) and early multi-linguals who had learned English between ages 2

and 6 years did not perform significantly different from each other [$F(1,48) = .24$, $p = .63$], but the infancy and early multi-linguals compared to the late multi-linguals who had learned English after the age of 6 years were significantly different [$F(1,78 = 3.90$, $p = .05$]. Adults performed significantly better than chance when they had learned English simultaneously with another language in infancy [$M=56\%$, $t(19) = 2.69$, $p <.02$] or learned English as a second language between the ages of 2 and 6 years [$M=57\%$, $t(29)=3.53$, $p <.02$] (Fig. 3.2). Those who learned English after the age of 6, however, did not perform significantly better than chance [$M=52\%$, $t(29)=0.82$, $p =.417$] (Fig. 3.2).

Follow-up analyses correlating years of experience [$r(67) = .16$, $p =.19$][4], self-rated proficiency [$r(79) = -.06$, $p = .60$] in English and minimal exposure to French [$r(79)= .09$, $p =.46$] had no influence on the results.
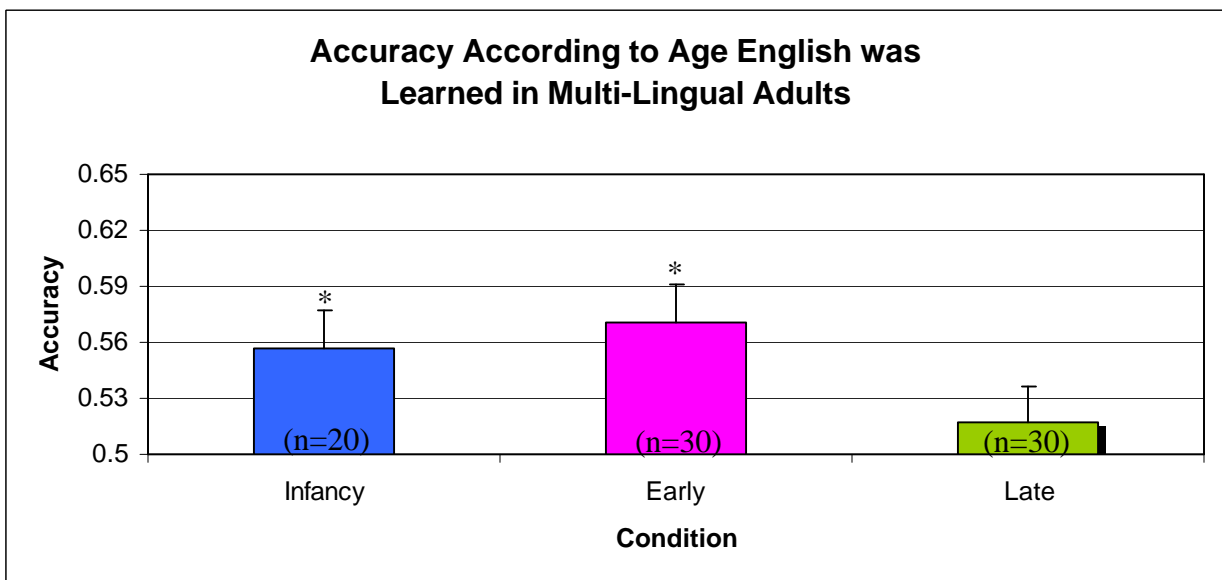
Figure 3.2. Mean accuracy performance of multi-lingual adults who had learned English: simultaneously with another language in infancy (Infancy), between age 2 and 6 (Early) and after the age of 6 (Late). The y axis represents mean accuracy and the x axis represents age at which English was learned. Error bars represent the standard error of the mean. *$p <.05$.

---

[4] Data for this analysis was only available for 68 of the 80 participants

3.4 Discussion

When tested on a visual language discrimination task, participants who had learned English as a second language after the age of 2 years (L2) performed more poorly than the participants who had learned English from birth (L1). Blocking sentences from the same speaker together (allowing the participant to develop familiarity with the speaker), versus a random speaker presentation, did not significantly improve performance for either group. However, the age at which a second language was learned did impact the adults' sensitivity to visual language cues. If English was learned simultaneously with another language in the infancy period (0-2 years) or early childhood (2-6 years), the participants performed significantly better than participants who had learned English as a second language after age 6 years. These results provide support for the hypothesis that an optimal period exists in early childhood in which exposure to visual speech information is necessary for maintaining sensitivity to visual language information. Interestingly, although changes are occurring during the first year of life with regard to infants' abilities to discriminate language visually (Weikum et al., 2007), the ability to discriminate languages visually in adulthood does not appear to rely on exposure starting in infancy. Indeed, exposure any time up to age 6 seems to provide the same protection for continued discrimination in adulthood. This suggests that an optimal period for visual language information extends, beyond the period of perceptual reorganization seen in infancy, into early childhood.

Optimal (or sensitive) periods have been previously identified for phonetic segment discrimination in auditory spoken languages, (for a review see Werker & Tees, 2005) and for acquisition of syntax in signed languages (Newport, 1990). The results from this study further

support these findings by showing that optimal periods exist for language discrimination, and on the basis of visual cues alone. Learning a new language when the optimal period for learning visual language information has passed may add to the difficulties encountered when acquiring a second language. Determining the visual cues that infants and adults use to discriminate languages visually will provide insight into the speech perception limitations faced by both first and second language learners.

3.5 Materials & Methods

3.5.1 Participants

Adults (120 participants) between the ages of 18 and 50 years (M=22.3 years) participated. Sixty had learned English from birth, and 60 had learned English as a second language after the age of 2 years. No participant was fluent in French. The other languages of the multi-lingual participants were quite varied, but the majority of the languages had an Asian origin.

3.5.2 Stimuli

Faces of three bilingual (French/English) speakers were recorded while they recited sentences in both English and French. See Appendix A for a list of the sentences. The sentences for each language were individually digitized to create 8 to 13 second video clips. The clips for each speaker were grouped in pairs. Each pair contained one speaker and two different sentence clips that were either in the same language (both English or both French) or different languages (English then French or French then English).

3.5.3 Procedure

      The adults each watched 48 clips presented as 24 pairs. Each pair of clips was played consecutively on a 17-inch computer monitor. All participants sat at eye level with the monitor, approximately 75 cm from the screen. Following the presentation of the second clip, the adults marked (pressed mouse buttons) to indicate whether they thought the clips were in the same language or different languages. In the random condition, the clips from all three speakers were presented randomly to each participant. In the blocked condition, 16 clips (eight pairs) from each individual speaker were presented consecutively before moving on to the eight pairs from the next speaker. The speaker order for the blocks was counterbalanced across participants.

3.6 References

Bernstein, L. E., Demorest, M. E., & Tucker, P. E. (2000). Speech perception without hearing. *Perception & Psychophysics*, *62*, 233-252.

Brutt-Griffler, J., & Varghese, M. (2004). Introduction. *International Journal of Bilingual Education and Bilingualism, 7*(2-3), 93–101.

Flege, J. E., & Fletcher, K. L. (1992). Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America, 91,* 370-389.

Flege, J. E., Munro, M. J., & MacKay, I. R.A. (1995). Effects of age of second language learning on the production of English consonants. *Speech Communication*, *16,* 1-26.

Hardison, D. M. (1996). Bimodal speech perception by native and nonnative speakers of English: factors influencing the McGurk effect. *Language Learning, 46*(1), 3-73.

Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context and talker variability. *Applied Psycholinguistics, 24*, 495–522.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 46-748.

Navarra, J., & Soto-Faraco, S. (2007). Hearing lips in a second language: Visual articulatory information enables the perception of L2 sounds. *Psychological Research, 71,* 4-12.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science, 14*, 11-28.

Pallier, C., Bosch, L., & Sebastián-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition, 64,* B9–B17.

Sebastián-Gallés, N., & Soto-Faraco, S. (1999). Online processing of native and nonnative phonemic contrasts in early bilinguals. *Cognition, 72*, 111–123.

Soto-Faraco, S., Navarra, J., Weikum, W. M., Vouloumanos, A., Sebastián-Gallés, N., & Werker, J. F. (2007). Discriminating languages by speechreading. *Perception and Psychophysics, 69,* 218.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America, 26*(2), 212-215.

Weikum, W. M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., & Werker, J. F. (2007). Visual Language Discrimination in Infancy. *Science, 316*, 1159.

Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology*, *46*(3), 233-251.

CHAPTER IV- Cues Infants Use to Discriminate Languages Visually[5]

4.1 Introduction

The world's languages can be divided according to many properties, but segmental and rhythmical differences are two salient properties that contribute to speech perception. Each language contains a characteristic set of phonetic segment differences that can be meaningfully combined to form different words. Although infants show broad-based discrimination to both native and non-native phonetic distinctions at birth, by the end of the first year of life, an infant's discrimination performance is maintained for speech sound segments that the infants continue to hear, but declines for the phonetic distinctions that are not phonemically contrastive in their native language (for reviews see Werker & Tees, 1999; Saffran, Werker & Werner, 2006). Conversely, discrimination performance on the specific subset of sound segments characteristic of an infant's native language improves (Kuhl et al., 2006).

Rhythmical properties also differentiate languages. Depending on where stress and emphasis are placed within words and/or sentences, languages can be roughly classified as having stress-timed (e.g. English), syllable-timed (e.g. French), or mora-timed (e.g. Japanese) rhythm (e.g. Ladefoged, 1975). Infants can auditorily discriminate their native language from a rhythmically unfamiliar language using low-pass filtered auditory speech, which removes much of the phonetic information, but maintains the rhythmical information (Mehler et al., 1988; Nazzi, Bertoncini, & Mehler, 1998).

Many caregiver-infant interactions involve face-to-face exchanges (e.g. Cohen & Beckwith, 1976), so a baby's experience with language is filled with talking faces. Since infants are raised in environments where often more than one person may be talking, it is of critical

---

[5] A version of this chapter will be submitted for publication. Weikum, W. M., Vatikiotis-Bateson, E., & Werker, J. F. Cues infants use for discriminating languages visually.

interest to determine the mechanisms infants use to attend to speakers in noisy situations. Studies with adults show that the ability to see a talker's face can dramatically improve speech perception in noisy situations (e.g. Sumby & Pollack, 1954). Additionally, the rhythmical movements and configurations of the talking face strongly correspond to the heard speech (Munhall & Vatikiotis-Bateson, 1998). Nonetheless, few studies have addressed the visual aspects of speech that adults and infants may use to attend to and perceive speech.

Recently, adults were shown to be capable of using solely visual speech information available in a speaker's face to discriminate languages, but they had to be familiar with one of the languages (Soto-Faraco et al., 2007). Visual cues alone are also sufficient for language (French and English) discrimination by infants as young as 4 months. However, infants who are only familiar with English experience a decline in their discrimination performance by 8 months, whereas babies exposed to both French and English continue to discriminate the languages at 8 months (Weikum et al., 2007). To better understand the factors that contribute to this developmental change in visual language discrimination, the following studies examined whether visual correlates of phonetic segments and rhythmical information are sufficient for visual language discrimination at 4 and 8 months. Determining the visual speech cues that infants use to tell languages apart will provide a better understanding of the visual speech information that infants are sensitive to in the speech signal, and might be using to learn language.

Visual correlates exist for both phonetic and rhythmical auditory language information. Typically, still pictures of the mouth or face producing a certain speech sound have been used to study the visual correlates of phonetic information (e.g. Massaro & Cohen, 1990; Montgomery & Jackson, 1983). However, still pictures do not capture the movement of the face as it is

pronouncing an auditorily heard speech sound. This paper therefore refers to the movement of

the lips, jaw, face and head when creating a discrete, short-term language sound (a segment), as

segmental visual phonetic or "visegmetic" information.

Infants' sensitivity to auditorily heard phonetic segments from unfamiliar languages

declines by the end of the first year of life (e.g. Werker & Tees, 1984), so we tested whether a

similar decline in sensitivity occurs for non-native visegmetic information. We hypothesized

that a decline in sensitivity to native vs. non-native visual phonetic segments may similarly

impair monolingual infants' visual language discrimination sensitivities at 8 months of age. The

first study examined whether infants use visegmetic information to discriminate languages

visually, and if insensitivity to the difference between native and non-native visegmetic

information impairs language discrimination at 8 months in monolingual infants.

Infants can auditorily discriminate rhythmically distinct languages when they are

presented forward, but not backward (Mehler et al., 1988). We therefore expected that infants

would perform similarly using visual rhythmical cues to discriminate languages. However, the

properties of reversed visual speech may be different from those in auditory speech. In contrast

to the odd sounds of reversed auditory speech, the reversed visual speech stream is very fluid and

natural in appearance. In fact, the reversed visual speech stream looks nearly identical to the

visual speech signal, with cues such as eye blinks and gasps for air as the only apparent oddities.

Although visual phonetic information and natural language rhythm are distorted in reversed

speech, it is possible that the rhythmical differences for each language (English and French)

remain systematically distinct when they are reversed. Ramus, Nespor, and Mehler (1999)

operationalized rhythmical differences between the languages by calculating the proportional

duration of vocalic and the variability of consonant intervals. Given that the movements of the

mouth, face and head are important in visual speech perception (Rosenblum & Saldaña, 1996) and can be used to reproduce the sound signal (Munhall & Vatikiotis-Bateson, 1998; Yehia, Rubin & Vatikiotis-Bateson, 1998), the vowel durations and consonant variability may also be available in visual speech. The dynamic duration of mouth openings may approximately correspond to vowels and the configuration of dynamic closures may roughly reveal the variability of consonants. Consequently, although the regular language rhythm conveyed through the kinetic movements of the face and mouth are disrupted through reversal, consistent rhythmical differences between the languages may nonetheless be maintained. If infants are sensitive to this reversed visual rhythmical language information, they may be capable of discriminating languages when they are played backward.

The visual forward and backward speech streams were first tested to ensure that they are visually discriminable. Evidence of discrimination for forward versus backward instances of the native language would be consistent with the hypothesis that the cues necessary for language discrimination are disrupted in the reversed speech. We expected that infants would be capable of telling the directions apart if there are sufficient differences between the forward and reversed versions of the visual speech stream, and thus likely different cues available in each direction.

The final study tested infants on their ability to tell the reversed instances of French and English apart. Given that infants auditorily fail to discriminate reversed languages (Mehler et al., 1988), one could predict that infants might similarly fail to tell languages apart visually when they are reversed. If however, the visegmetic cues are disrupted by playing the speech backward, but consistent visual rhythmical differences are nonetheless maintained, we expect that infants may succeed in telling apart the reversed visual speech stream.

This research therefore examines the phonetic and rhythmical cues that may be sufficient for discriminating English and French in infancy. Infants from monolingual English homes were tested at the youngest age that they have been shown to discriminate languages visually (4 months), and at the age when a decline in visual language discrimination occurs (8 months). Understanding which visual speech cues facilitate and/or impair language discrimination at both 4 and 8 months in monolingual infants will significantly enhance our understanding of the developmental trajectory of visual language perception, as well as the visual speech cues available to infants. Determining the visual speech cues that infants are sensitive to is important for understanding how infants perceive the visual speech signal, and provides a platform for further studies investigating how infants use visual speech to help them learn language.

4.2 Experiment I

4.2.1 Method

This experiment tested whether visegmetic information is sufficient for language discrimination at both 4 and 8 months. In order to isolate the visegmetic information, it was necessary to remove the rhythmical information as a consistent cue differentiating the languages. Dehaene-Lambertz and Houston (1998) distorted the rhythmical information in auditory speech sentences by rearranging the words to create scrambled sentences. We therefore manipulated the visual speech stream by cutting sections to isolate visual phonetic segments. The sections were cut at 200 msec because it is below the average length of the syllables used in our study clips (English average, 236 msec; French average, 223 msec). Cutting our segments just below the average length of the syllables in our sentence stimuli ensured that the segments were long enough to capture visegmetic information, but disrupted the influence of syllable or sentence-

level rhythm. The short visegmetic clips were then randomly reordered for each sentence to create a "scrambled" video clip. The scrambled clips maintained some of the segmental visual phonetic information, but distorted the rhythmical language information because the continuous rhythmical movement of the face had been reordered.

Infants were habituated to the scrambled video clips of bilingual French/English speakers reciting sentences in one of the languages. Infants were then shown the opposite language or the same language to see if their looking times changed. If infants noticed the language change, we expected them to look longer at the test trials when the language changed, compared to when it stayed the same, indicating that the infants noticed a change in the language.

4.2.2 Results

A repeated-measures analysis of variance (ANOVA) analyzing age (4 vs. 8 months), condition (same vs. switch) and trial (final habituation vs. test trials) revealed a significant main effect for age [$F(1,44) = 14.08$, $p < .01$] and condition [$F(1,44) = 6.04$ ($p < .02$)]. Post-hoc simple main effects analyses showed that there was a significant difference between the same-language test trials and the different-language test trials at 4 months [$F(1,22) = 4.92$, $p < .05$], but not 8 months ($F(1,22) = .48$, $p = .50$; Fig.4.1]. At 4 months, the infants can rely on visegmetic cues to discriminate languages, but at 8 months the visegmetic cues alone are not sufficient for language discrimination.
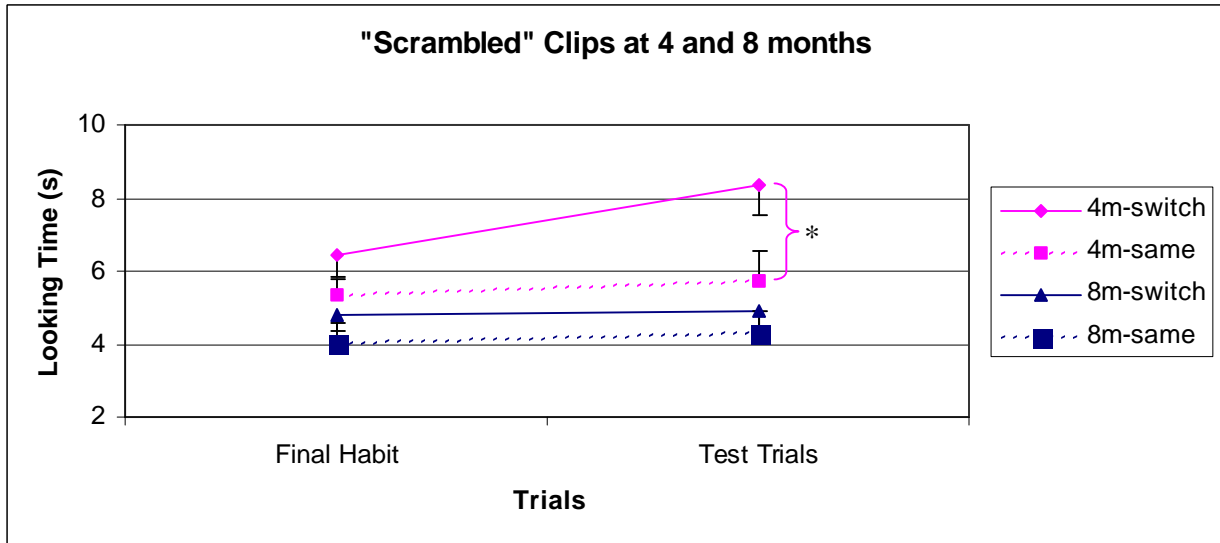
Figure 4.1. Mean looking time in seconds to silent talking faces that had been scrambled to preserve visual phonetic cues and disrupt rhythmical cues. The y axis represents infant looking time; the x axis represents the trial that the infant was shown (final habituation trials or test trials). Error bars represent the standard error of the mean. Experimental (language switch) and control (language same) conditions for infants at 4 and 8 months, *p < .05.

4.3 Experiment II

4.3.1 Method

Infants were tested to determine whether visual speech cues are perceptibly different/distorted in the reversed speech stream. We tested the infants at 4 and 8 months to see if they could tell apart the forward and backward visual speech clips from their native language English. The infants were habituated to the English clips played in one direction, and then shown different English clips in the opposite direction.

4.3.2 Results

Infant discrimination of backward visual speech versus forward visual speech was analyzed using a repeated-measures ANOVA. Age (4 vs. 8 months), condition (same vs. switch) and trial (final three habituation vs. test trials) were analyzed and revealed a significant main effect for age [$F(1,44) = 4.5, p < .05$], significant trial by age interaction [$F(1,44) = 4.3, p < .05$] and significant trial by condition interaction [$F(1,44) = 6.88, (p < .02)$], but no age by condition interaction [$F(1,44) = .15, p = .70$]. Simple main effects follow-ups showed that the infants looked significantly longer during the test trials when direction of the language switched, compared to the trials when the direction stayed the same at both 4 months [$F(1,22) = 5.61 , p < .05$] and 8 months [$F(1,22) = 4.4, p < .05$; Fig. 4.2].

Figure 4.2. Mean looking time in seconds to silent talking faces that were played in different directions. The y axis represents infant looking time; the x axis represents the trial that the infant was shown (final habituation trials or test trials). Error bars represent the standard error of the mean. Experimental (direction switch) and control (same direction) conditions for infants at 4 and 8 months, *$p < .05$.
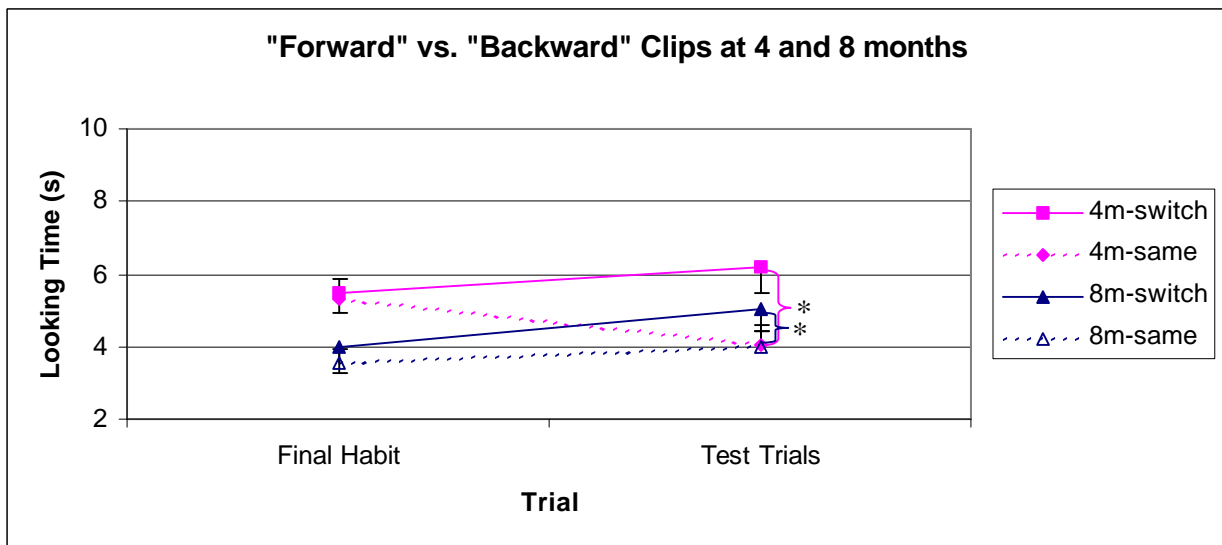
4.4 Experiment III

4.4.1 Method

Infants were tested on their discrimination of the reversed visual English and French clips. Infants at 4 and 8 months were habituated to reversed clips from one of the languages, and then shown reversed clips from either the opposite language or the same language for six test trials. If the infants perform similarly on the visual task, compared to how infants performed on the auditory speech task, we expected them to fail to tell apart the languages. However, if, as described earlier, the properties of the visual signal preserve rhythmical information in the reversed form, (potentially facilitating language discrimination) we expected language discrimination.

4.4.2 Results

A repeated-measures ANOVA examining age (4 months vs. 8 months), condition (language switch vs. language same) and trial (habituation vs. test) revealed significant main effects for trial [$F(1,44) = 4.43$, $p < .05$], age [$F(1,44) = 10.48$, $p < .01$] and condition [$F(1,44) = 4.65$, $p < .05$] as well as a significant interaction for trial by condition [$F(1,44) = 18.75$, ($p < .01$], but no age by condition interaction [$F(1,44) = .002$, $p = .96$]. Simple main effects analyses showed that the infants looked significantly longer at the backward test trials when the language switched compared to the trials when the language stayed the same at 4 months [$F(1,22) = 4.475$, $p < .05$] and 8 months [$F(1,22) = 7.9$, $p < .02$; Fig. 4.3]. Infants at both ages can tell languages apart when they are viewed in reverse.

"Backward" Clips at 4 and 8 months

Legend:
- 4m-switch
- 4m-same
- 8m-switch
- 8m-same

Figure 4.3. Mean looking time in seconds to silent talking faces that were reversed to disrupt visual phonetic and rhythmical language cues. The y axis represents infant looking time; the x axis represents the trial that the infant was shown (final habituation trials or test trials). Error bars represent the standard error of the mean. Experimental (language switch) and control (language same) conditions for infants at 4 and 8 months. *$p < .05$.
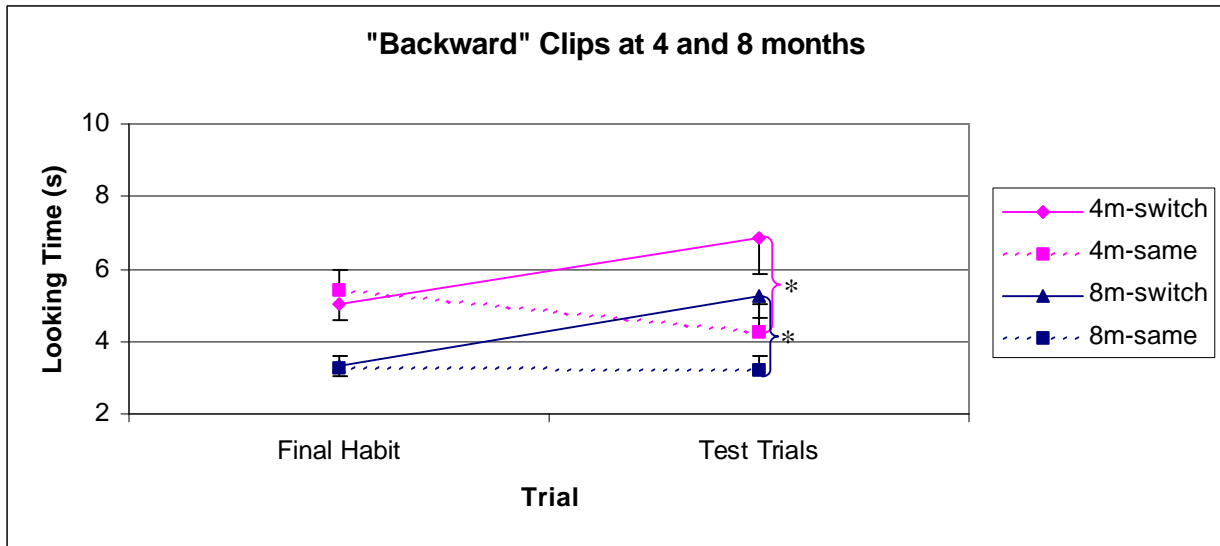
4.5 Discussion

At 4 months of age infants are able to discriminate the languages visually when the rhythmical information is distorted and some segmental visual phonetic information is preserved. At 8 months of age, however, monolingual English infants again fail on the visual task suggesting that there is reduced sensitivity to the difference between native and non-native visegmetic information that causes monolingual infants to fail to notice the language switch. In essence, the infants fail to notice the language switch because they fail to notice a change in the segmental visual phonetic information.

The infants perceived the forward and backward versions of the visual speech clips as different, suggesting that there is a critical difference between the clips, and that different cues may thus be available for language discrimination. This experiment was also conducted as a preliminary step to help determine whether or not segmental visual phonetic cues are preserved in the reversed speech. If infants are only paying attention to the shapes made by the face, for instance the amount of tongue protrusion and lip rounding, they may not notice a difference between the directions of the video clips. If, however, infants are sensitive to the language information related to the actual formation of the visual phonetic units, and/or to the sentence-level rhythm of the language, then the infants would be able to tell apart the forward and backward clips. We found that the infants are sensitive to the direction of the language, so support is provided for the proposition that critical cues related to language discrimination are at least different, and likely distorted, in the reversed speech stream.

The reversed manipulation experiment examined whether visual speech information patterns similarly to auditory speech information, or if there is something distinctive about the visual speech signal that is conserved when it is reversed. Contrary to auditory speech findings, the monolingual English infants successfully discriminate the reversed languages at 4 months. More surprisingly, they also succeed at 8 months, even though previous findings (Weikum et al., 2007) show that monolingual English infants fail to discriminate languages when the visual speech signal is presented forward.

The discrimination success of infants at both 4 and 8 months with the backward visual speech suggests that reversed visual speech may not share the same disruptions to properties in the speech signal as reversed auditory speech. Time-critical and ordinal segmental properties are distorted in the reversed auditory speech signal (Sheffert, Pisoni, Fellowes & Remez, 2002), and

physiologically, many of the sounds of reversed speech cannot be produced by the human vocal tract. Thus, the unnatural attributes of the reversed auditory speech signal may be impairing infants' performance on auditory language discrimination. Studies have shown hemispheric specialization for natural forward speech, but not for backward speech in infants (Peña et al., 2003), and infants prefer to listen to speech over complex non-speech sounds (Vouloumanos & Werker, 2007). Perhaps, for visual speech, the reversed stimuli are more biologically plausible and thus subject to different processing mechanisms. Although the phonetic segments are likely less identifiable in visual speech, the consistent differences in visual language rhythm of reversed visual speech may be conserved and facilitate language discrimination in our task.

Monolingual infants may therefore fail to discriminate languages visually at 8 months when the speech is presented forward, and succeed when it is presented backward, because the infants are sensitive to and respond differently to the cues available in each direction. The infants may perceive the backward and forward speech as different, because the normal forward speech activates processing mechanisms specific to language (such as the motions related to segmental visual phonetic information). The reversed speech may significantly degrade segmental visual phonetic cues, and thus be subject to lower-level processing (perhaps non-linguistic), that relies solely on the plausible rhythmical differences maintained in the reversed clips.

Imaging studies will help determine how infants are processing the reversed visual speech. Near Infrared Spectroscopy (NIRS) shines light through an infant's scalp to measure oxygenation levels related to blood flow. Studies have revealed hemispheric specialization related to language tasks (Peña et al., 2003). It would therefore be plausible to examine both forward and backward visual speech to determine if hemispheric specialization is evident in one

speech direction (forward) but not the other (backward). An absence of hemispheric specialization during the backward speech task would support the view that a lower-level or different processing mechanism may account for the success of infants on the discrimination of backward languages at both 4 and 8 months.

Behavioural studies will also be conducted to confirm the conditions under which languages can be discriminated visually by infants. For instance, infants should be tested on rhythmically dissimilar languages that they are unfamiliar with, to see if infants can rely on rhythmical differences to tell the languages apart (i.e. test babies from Japan on French and English). Conversely, infants can be tested using languages that are rhythmically very similar (e.g. Spanish and Catalan) to help determine when and if solely segmental visual phonetic information can be used to discriminate languages visually.

4.6 Materials & Methods

4.6.1 Participants

Data from 132 monolingual English hearing infants were included in the final sample. Half the infants were 4m (M=4 months 19 days, SD=16 days) and half the infants were 8m (M=8 months 18 days, SD=16 days). There were 12 infants for each condition at both ages, with an equal distribution of males and females. Additional infants were tested and excluded from the analyses for failure to habituate (10), fussiness (25), disinterest/failure to watch the screen (2), parental interference (6), experimenter error (5), or statistical outlier (1). For the second study (backward vs. forward visual speech) data for the control condition were taken from Weikum et al., (2007). The data from all six infants who watched only English clips and were tested in the control condition were used again for this study.

4.6.2 Stimuli

The stimuli used were from a previous experiment (Weikum et al. 2007). Bilingual (French/English) speakers were filmed while they read sentences from a children's story in both languages (See Appendix A for a list of the sentences used). The face of each speaker reciting a sentence was digitized into *silent* 8 to 13 second video clips.

For the scrambled stimuli, we did not use still faces to represent the visual phonetic information as has been done in previous studies. For example, Montgomery and Jackson (1983) used the static features of tongue height and degree of lip rounding or spreading to identify vowels and Massaro and Cohen (1990) suggested that the degree of opening represented by the lips is important for distinguishing between certain consonant sounds. We reasoned that it is imperative to provide information about the movement of the face as it creates visual phonetic segments, as this is the most biologically plausible and informative source. Although it would have been ideal to divide the visual speech stream into individual "visual phonetic segments" that represented the phonetic segments of each language, there is a strong influence of coarticulation and indeterminacy as to where to visually capture the "start" and "end" of each visual phonetic segment. It was therefore necessary to select a method that allowed us to randomly choose dynamic segmental visual phonetic cues. Each clip was systematically cut into segments that captured visegmetic information and then randomly reordered to disrupt rhythmical information. In order to prevent syllable rhythm from potentially influencing the results, the visual speech stream needed to be cut into segments that were below the length of the average syllable length in our test sentences. To calculate the average size of the syllable segments for each language, the auditory speech stream for each sentence was edited to remove

pauses. The total duration of speech for each sentence was then divided by the respective number of syllables for each sentence. The English sentences used in the study averaged 236 msec and the French sentences averaged 223 msec. The clips were therefore cut into 200 msec segments because it is below the length of the average syllable length in our test sentences, thereby diminishing syllabic rhythm as a possible influencing factor.

In the second experiment the English visual speech stimuli from Weikum et al. (2007) were shown to the infants normally (forward) and played in reverse (backward). The third experiment used the French and English stimuli from Weikum et al. (2007) and all the video clips were played in reverse.

4.6.3 Procedure

In all three experiments, the infants were seated on their parent's lap in a dimly lit, sound-attenuated room approximately 4 feet from a 27" television screen. The parents were instructed to remain silent, refrain from directing their infant's attention towards the television screen at any time, and wore blackened sunglasses to prevent them from seeing the video clips and potentially influencing their infant. The faces of the three bilingual speakers silently reciting sentences were approximately life-sized and displayed on the television screen sequentially.

For the first experiment, the infants were habituated to the scrambled clips from one of the languages, and then shown the opposite language for six test trials (test condition), or six trials from the same language (control condition). The test trials for both conditions contained three new sentences from each speaker, and were repeated once to give the infants time to notice the subtle language change. The speakers were always shown in the same order for each infant,

even across the habituation and test trials (counterbalanced between infants), and the infants were always presented with different sentences.

Infants were tested using the same procedure in the second (forward vs. backward English) experiment. Half the infants were habituated to the reversed clips from English and half were habituated to the forward clips from English. Immediately following habituation, half the infants viewed six test trial clips played in the opposite direction (test condition) or half viewed six test trials from the same direction (control condition).

The procedure for the third experiment (backward English and French) was identical to that used in the two previous studies, except that the clips were all reversed versions of the sentences. Half the infants were habituated to the reversed clips from English and half were habituated to reversed French clips. Following habituation, half the infants were shown reversed clips from the opposite language for six test trials (test condition), and half were shown six test trials from the same language (control condition).

4.7 References

Cohen, S. E., & Beckwith, L. (1976). Maternal language in infancy. *Developmental Psychology, 12,* 371-372.

Dehaene-Lambertz, G., & Houston, D. (1998). Language discrimination response latencies in two-month-old infants. *Language and Speech, 41*, 21-43.

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), F1–F9.

Ladefoged, P. (1975) *A Course in Phonetics*. New York: Harcourt Brace Jovanovich.

Massaro, D. W., & Cohen, M. M., (1990) Perception of synthesized audible and visible speech. *Psychological Science, 1*(1), 55–63.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29*, 143-178.

Montgomery, A. A., & Jackson, P. L. (1983). Physical characteristics of the lips underlying vowel lipreading performance. *Journal of the Acoustical Society of America, 73*, 2134-2144.

Munhall, K. G., & Vatikiotis-Bateson, E. (1998). The moving face during speech communication. In R. Campbell, B. Dodd & D. Burnham (eds.) *Hearing by Eye II: Advances in the Psychology of Speechreading and Auditory-visual Speech* (pp. 123-139). East Sussex, UK: Psychology Press Ltd.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception & Performance, 24*(3), 756-766.

Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., & Mehler, J. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences of the United States of America, 100*(20), 11702-11705.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*(3), 265-292.

Rosenblum, L. D., & Saldaña, H. M. (1996). An audiovisual test of kinematic primitives for visual speech perception. *Journal of Experimental Psychology: Human Perception and Performance. 22*(2), 318-331.

Saffran, J. R., Werker, J. F., & Werner, L. A. (2006). The infant's auditory world: Hearing, speech, and the beginnings of language. In R. Siegler & D. Kuhn (Eds.), *Handbook of child development* (Vol. 6, pp. 58-108). New York: Wiley.

Sheffert, S. M., Pisoni, D. B., Fellowes, J. M., & Remez, R. E. (2002). Learning to recognize talkers from natural, sinewave, and reversed speech samples. *Journal of Experimental Psychology, 28*(6), 1447-1469.

Soto-Faraco, S., Navarra, J., Weikum, W. M., Vouloumanos, A., Sebastián-Gallés, N., & Werker, J. F. (2007). Discriminating languages by speechreading. *Perception and Psychophysics, 69,* 218-231.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America, 26*(2), 212-215.

Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science, 10*(2), 159-164.

Weikum, W. M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., & Werker, J. F. (2007). Visual Language Discrimination in Infancy. *Science, 316,* 1159.

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

Werker, J. F., & Tees R. C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology, 50*, 509-535.

Yehia, H. C., Rubin, P. E., & Vatikiotis-Bateson, E. (1998). Quantitative association of vocal-tract and facial behavior. *Speech Communication, 26*, 23-43.

CHAPTER V – General Discussion

5.1 Introduction

Visual speech information may play a more important role in language acquisition than previously anticipated. By highlighting the importance of visual speech information for language discrimination, this research demonstrates that facial information may play an integral role in preparing both monolingual and multi-lingual infants to adapt to their language-learning environments. From a very early age, infants are prepared to separate languages on the basis of visual cues, and advantageously maintain the ability to separate languages if it is required in their language-learning environment. These changes in sensitivity to visual language information suggest that there is a privileged period in early life during which exposure to visual speech has the most impact. The adult study from chapter 3 tested whether exposure to English in infancy (0-2 years), early childhood (age 2-6 years) or late childhood (6+ years) impacts the ability to discriminate languages visually as an adult. Although there is no advantage to learning English in infancy compared to early childhood, a lack of experience with the visual concomitants of language before age 6 has a lasting impact on the adults' abilities to discriminate English from an unfamiliar language. As the final sets of infant studies show, segmental visual phonetic information may become selectively tuned according to an infant's native language across the first year of life, and result in reduced sensitivity to non-native visual language information. This reduced sensitivity to visual speech information may in turn add to the difficulties encountered when children and adults are learning a second language.

5.2 Conclusions

The ability of young infants to discriminate languages using solely visual speech information gives insight into how the brain may be organized to learn language. The research discussed in this thesis shows that babies as young as 4 months can tell languages apart visually. This shows that, before language is even understood, infants are sensitive to visual cues that differentiate languages and use these cues to tell languages apart. This finding supports a preparedness view of perception where the brain appears to be equipped to categorize and distinguish consistent patterns, even at the level of a talking face.

This thesis research also shows that if the patterns available for telling languages apart prove to be consistent in an infant's environment (i.e., the infant is regularly exposed to more than one language) infants will continue to discriminate or maintain a level of sensitivity necessary for separating the languages. This was shown quite clearly by the studies indicating that infants, at 8 months, only succeed in discriminating two languages if they have had consistent exposure to both languages. This research therefore supports the view that infants come prepared for perceiving more than one language, and that subsequent (re)organization occurs to match the language-learning environment.

These findings also support and provide insight into how adaptable the brain is to learning more than one language. Young bilingual infants can use solely visual speech information to separate the languages they are learning, and maintain this ability even when their monolingual English counterparts fail. This helps us better understand how bilinguals may be capable of separating their languages as they are learning them. This also suggests that infants are grouping their languages as they learn them. For instance, the visual findings suggest that infants do not necessarily just learn one large language and then later determine that they are

learning two languages, and group them accordingly. Instead, the infants' abilities to visually separate languages may help to keep their languages separate while they are learning them.

This preparedness in early infancy and subsequent (re)organization of visual language discrimination abilities coincides with what appears to be an optimal period for maintaining sensitivity to visual speech cues that permit language discrimination. Although adults can tell a language they are familiar with (English) from another rhythmically distinct language, the familiar language must have been learned before the age of 6 years. The adult data support the view that there is a period in early childhood or infancy that may lay the foundation for later visual language perception abilities.

Finally, the visual speech cues to which infants may be sensitive were examined. Both rhythmical language and segmental visual phonetic information (alone or in combination) can be used to discriminate languages visually at 4 months. Visegmetic information is not sufficient to allow language discrimination at 8 months if infants are only familiar with one of the languages. The decline in sensitivity to non-native visual segmental information closely resembles the auditory speech literature, which shows a decline in sensitivity to non-native phonetic contrasts near the end of the first year of life (Werker & Tees, 1984). It appears to be the case that visual speech discrimination follows a trajectory that is very similar to auditory speech discrimination. This is consistent with theories suggesting that speech perception may rely on amodal information available (to a degree) in both visual and auditory speech information.

Overall, these findings reveal a remarkable readiness in early infancy for perceiving the cues that distinguish linguistic communities, and show how the tight coupling between experience and perceptual development leads to different, but aptly beneficial outcomes for both bilinguals and monolinguals.

5.3 Considerations

Given that infants are capable of discriminating their native language from an unfamiliar language, it was expected that infants might prefer watching their native language. However, no strong claims to preference can be made on the basis of the preference study described in the introduction. Furthermore, the rate of habituation data, and amount of looking on the first trials in the infant discrimination studies show no significant difference between the French and English clips. However, since the trial duration was only 16 seconds in the habituation study, and 20 seconds in the preference study, it may be the case that the trial length is not long enough to pull apart a preferential difference in the habituation study.

At 8 months, monolingual English infants failed to notice a switch in the visual language discrimination task. It could be the case that the failure of the monolingual infants was due to the fact that monolingual infants at this age are not interested or capable on this task. We eliminated this possibility in the same study (Chapter II), because bilingual infants at 8 months could in fact discriminate the languages. Moreover, the second infant study (Chapter IV) shows quite clearly that monolingual infants at 8 months of age can succeed on a visual language task, because they are capable of telling the languages apart when they are presented in reverse. Thus, it is likely that the failures to tell the languages apart in the normal forwards condition and scrambled condition in 8 month monolingual infants do not represent an inability to complete the task, but rather an inability to tell the languages apart.

The decline in visual language information at 8 months in monolingual infants has been likened to the similar decline in non-native auditory speech perception. Although the tasks are dissimilar, the effects may have a common origin. Because less information is communicated through visual speech, compared to auditory speech (e.g. Erber, 1974), it could be the case that a

reduced sensitivity to visual language information from a non-native language allows the visual speech information from the unfamiliar language to somehow become pulled into the infant's native language category. The effect may be similar to the Native Language Magnet Model proposed by Kuhl (1991; 1993), which suggests that poor instances of a phonetic category are pulled into and categorized according to the closest category prototype that they resemble. Thus, as monolingual infants' sensitivity to the non-native visual attributes of French declines, perceptually they may fail to notice the visual difference between French and English because they are collapsing all the segmental visual phonetic information into English categories.

However, if the infants are still capable of discriminating the languages on the basis of rhythmical cues (as the backward visual speech data suggests), it remains unclear why the 8-month-old infants were not able to tell apart the languages using the rhythmical cues when the clips were presented forward. Within the auditory speech literature, at different periods during development, infants are more sensitive and responsive to different cues in the speech signal. For instance, at the word level, Saffran, Aslin and Newport (1996) showed that 8-month-old infants can segment a continuous speech stream into words based solely on the statistical relationships between the neighboring speech sounds. However, other studies have shown that certain speech cues may influence or even override this statistical segmentation. For example, Johnson and Jusczyk (2001) showed that 8-month-olds weigh speech cues, such as stress patterns, more heavily than the statistical cues, and Thiessen and Saffran (2003) found that the use of stress cues and statistics changes depending on the age of the infant. Thus, there may be a similar visual effect at 8 months, where infants are paying particular attention to the segmental visual phonetic information, and the change in rhythm is not enough to be detected by their

speech perception system at that time. Additionally, when the clips are reversed visually, the visual rhythm differences may be more dramatic, facilitating discrimination.

The success of the bilingual infants at 8 months in the visual language discrimination task is being interpreted as a maintenance advantage conferred by exposure to both languages. However, within the auditory speech literature, differences and delays have been documented for bilingual learning infants on certain tasks (e.g. Bosch & Sebastián-Gallés, 2003; Fennell, Byers-Heinlein, & Werker, 2007). However, studies have failed to find differences in the tuning of phonetic perception (e.g. Burns, Yoshida, Hill & Werker, 2007). In fact, studies specifically related to discriminating languages auditorily have failed to find differences between monolingual and bilingual infants (Bosch & Sebastián-Gallés, 2001). Is it the case then, that the bilingual success at 8 months is merely a delay in speech perception maturation? While plausible, this explanation is also unlikely given the results from the adult data on visual language discrimination. First, adults who had no exposure to visual speech information from their second language in early childhood perform at chance on the language discrimination task. This strongly suggests that early experience with visual speech information is essential for setting and maintaining a foundation for visual language discrimination abilities. Secondly, Soto-Faraco et al. (2007) showed that adult participants who knew both languages they were tested on in a language discrimination task performed better than participants who were only familiar with one of the languages. Presumably, if experience with both languages affords an advantage in adulthood, it is not inconceivable that a similar advantage may therefore also exist in infancy. Regardless, follow-up studies with older (10 month French/English bilingual) infants are currently being conducted in order to test this alternative explanation.

Although it was discovered that infants who are only familiar with English fail on the discrimination task by 8 months, adults who are only familiar with English succeed when English was learned before age 6. The reasons for the success of adults familiar with one of the languages and apparent failure in 8-month-old infants may be two-fold. It could be the case that adults regain some lost sensitivity through extended experience with the language or through a more strategic application of attention during the task. However, given that adults who learned English after the age of 6 fail regardless of self-rated proficiency or years of experience with the language, it is probably more likely that the adult task is just a more sensitive measure of discrimination. Since infants cannot verbally describe whether or not they detect a difference, the more indirect methods of data collection that must be used with infants are likely less sensitive.

5.4 Future Studies

Future studies should continue to address how visual speech information is more than just a redundant speech signal and may in fact, as part of an integrated percept, provide a more clear, reliable and learnable speech signal. Instead of just being helpful in noisy situations, it will be important to investigate visual speech from the contribution it has when learning a language (whether it be an infant learning their first language, or a child or adult learning their second language).

The bilingual infant studies presented in this thesis only address bilingual infants who know both languages being tested. It is of critical interest to determine whether the ability to separate languages visually is generalizable across different languages, or specific to the languages being learned. This will give valuable insight into how the brain organizes visual

language information in infancy. Insensitivity to segmental visual phonetic information appears

to be responsible for the failure of 8-month-old monolingual infants, so it may be expected that

the ability to tell the languages apart is solely derived from experience with the languages in

question. If, however, the language discrimination abilities of bilingual infants represent a more

global, and generalizable neural process, it might be expected that exposure to multiple

languages facilitates the ability to visually discriminate all languages, or perhaps languages

within a certain rhythmical category. Infants who are learning English and another language (not

French) could be tested on the French/English visual discrimination tasks. The other languages

the infants are learning would then be analyzed to determine whether the rhythm of the other

language matches one of the languages they are being tested on, or is different. This will help

narrow down whether solely rhythmical information is sufficient for language discrimination,

and also determine whether rhythmical language categories are an influencing factor in bilingual

visual language discrimination. Generally, this study would show whether the ability of

bilingual infants to maintain their visual language discrimination abilities extends to all other

languages, other languages within the rhythmical language categories they have been exposed to,

or just the languages that they are familiar with.

The finding that 8-month-old monolingual infants can still discriminate French and

English if played in reverse suggests that strictly rhythmical information can be used to tell the

languages apart. Ramus, Nespor, and Mehler (1999) proposed that the vowel durations and

consonant variability of the words from a language result in the specific rhythmical differences

that have been found to characterize languages. Using data from the existing literature on

rhythmical speech discrimination, Ramus and colleagues compared eight different languages and

calculated how discriminable they are given their differences in vowel duration and consonant

variability. Interestingly, the findings suggest that, using only rhythmical information, French and English should be discriminable at about 60%, which is strikingly similar to the discrimination results of our study where adults watch pairs of silent video clips and decide whether the French/English bilinguals are reciting sentences in the same or different languages. Thus, analyzing visual speech characteristics corresponding to vowel openings versus consonant closures may reveal whether visual information can be used as another rhythmical or prosodic cue to facilitate language discrimination and perhaps even prosodic bootstrapping.

The role of rhythmicity can be further tested by designing studies using artificial talking heads. The auditory speech signal could be manipulated so that all the consonants and vowels are replaced with token consonants and vowels, similar to the manipulations done by Ramus and Mehler (1999) with auditory speech. To further characterize the ability of infants to discriminate languages using only rhythmical information, Ramus and Mehler (1999) re-synthesized the auditory speech from real utterances so that the vowels were replaced with an "A" and the consonants were replaced with an "S", "L", "T", "N", or "J" depending on their manner of articulation. This stringent control preserves only the alternation of vowels and consonants, as well as the rhythm and intonation. Amazingly, infants can still discriminate the languages. Ramus (2002) also found that French newborns can discriminate between Dutch (stress-timed) and Japanese (mora-timed), when the intonational cues are removed, but the rhythmical information is left intact. It would be fascinating to create an artificial talking head with facial gestures that maintain the vowel and consonant durations characteristic of the languages, but only make the facial movements characteristic of the vowels and consonants described above. The artificial talking heads could then recite the simplified language sentences. Thus, only the rhythm of the opening and closing of the mouth would change to represent the duration of

vowels and consonants represented in each sentence. If infants are sensitive to visual rhythmical

information, they should be able to tell apart the languages using the rhythmical movements of

the face as the cue.

Hollich, Newman, and Jusczyk (2005) have shown that infants can use a single visual

waveform, representing the auditory amplitude of the speech signal, to parse the speech stream.

It would therefore be interesting to test whether infants can use visual waveforms, representing

French and English sentences, to discriminate languages. This would definitely confirm whether

visual language rhythmicity also provides enough information for discriminating between

languages or identifying speakers of one's native language.

If rhythmical information permits infants to tell apart the languages, testing newborns to

see whether they can tell languages apart visually would provide evidence for an early

preparedness for language discrimination as well as support for speech as a modality-neutral

signal. Infants could be tested while watching silent faces talking. Although newborn vision is

far from accurate, the gross motions of the face may be enough to convey the rhythm of the

language. Indeed, Munhall, Kroos, Jozan, and Vatikiotis-Bateson (2004) have shown that a

limited range of visual-spatial frequency spectrum (e.g. low-pass filtered dynamic facial stimuli)

is sufficient for audio-visual speech processing in adults. When a newborn is watching sentences

from their native vs. an unfamiliar language, it might be expected that there are differences in

heart rate, in sucking amplitude on a blind nipple attached to a transducer or in neural activation

found using Near Infrared Spectroscopy (NIRS). NIRS uses points of light shone through an

infant's scalp to measure oxygenation levels related to blood flow. Testing infants with NIRS

would not only help determine whether infants respond differently to their native language

versus an unfamiliar language when it is presented visually, but evidence of hemispheric

specialization when processing visual speech would strongly support the view that infants are born prepared to process language multi-modally.

Future studies should also investigate the role of the integrated speech percept, to determine if a combination of modalities creates an enhanced effect that is better than either modality on its own. Research with adults suggests that combined auditory and visual cues in second language learners leads to a stronger perceptual effect than can be seen under solely auditory or visual circumstances alone (Navarra & Soto-Faraco, 2007). Understanding when and how this integrated percept develops will help determine how the brain is predisposed to learn language. Furthermore, a better understanding of the advantages (sensitivities) that infants have for processing different types of language information will give a much clearer understanding as to what may impair or hinder second language acquisition, and perhaps even provide insight into strategies that second language learners may use to help them acquire languages more easily.

After obtaining a better understanding of how an integrated percept may benefit language acquisition in adults, it will be important to further systematically investigate audio-visual language integration in infancy. Auditory language discrimination studies have only investigated infants' abilities to discriminate languages up to 4 months (Nazzi, Jusczyk, & Johnson, 2000). Language discrimination using both normal and low-pass filtered auditory speech should therefore be tested on infants older than 6 months. The results can then be compared to infants' discrimination abilities when both auditory and visual information is available. Technology such as NIRS with infants may provide a valuable tool for comparing the neural responses of both auditory and visual information presented alone and as an integrated percept. For strictly visual language information, testing with NIRS will help determine which areas or hemispheres of the brain are activated by solely visual language information. This can then be compared to the

activation created by strictly auditory language information to see if similar areas are activated to the same information presented by different modalities. The amount of activation created by the additive values of the auditory + visual speech stimuli can then be compared to the activation generated by an integrated auditory/visual speech stimuli. Comparing the separated and integrated neural responses will help us understand if the brain, in infancy, is designed to process integrated speech percepts. Testing at different infant ages would further help determine whether the brain processes both signals separately and then develops an integrated percept through experience, or is designed to integrate initially. For instance, if the integrated speech percept results in reduced activation compared to the combination of solely auditory + solely visual speech, then it suggests that the brain is initially wired to perceive speech as an integrated percept. If however, the amount of activation created by the integrated percept diminishes as the infant ages, but the amount of activation to solely visual and auditory speech information stays the same, this is more supportive of a trajectory whereby visual and auditory information become integrated through experience with both. These studies will give a clearer indication as to whether the language system is designed to be multi-modal from the outset, or whether visual information is more of a redundant signal that is developed in conjunction with auditory cues and most helpful under noisy circumstances. Research in audio-visual language acquisition will therefore help determine whether speech functions better in a multimodal environment, and essentially address whether the brain is designed to process speech more efficiently as a multimodal signal.

Overall, the findings and proposed research studies described in this thesis are designed to help us understand how the infant processes visual language information (uni-modally or multi-modally) from birth and throughout development. Understanding the changes in neural

sensitivity to both auditory and visual language information, as they are related to language experience, provides insight into the methods by which the brain (re)organizes to accommodate input from the environment. Determining the properties of speech (e.g. phonetic and rhythmical) that are susceptible to neural (re)organization reveals the types of stimuli that are sensitive to optimal periods during development. This will provide important insight into how neural development is shaped by language experience, and may be used to set a critical baseline to which abnormal speech or brain development may be compared. Thus, understanding the critical factors mediating the multi-modal perception of speech and the changes in sensitivity to speech information during development is important for understanding how the brain processes information to accommodate multi-modal environmental influences.

5.5 References

Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy, 2*, 29-49.

Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language specific vowel contrast in the first year of life. *Language and Speech*, *46*, 217-244.

Burns, T. C., Yoshida, K. A., Hill, K., & Werker, J. F. (2007). The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics, 28*(3), 455-474.

Erber, N. P. (1974). Discussion: Lipreading skills. In R. E. Stark (Ed.), *Sensory capabilities of hearing impaired children* (pp. 69-73). Baltimore, MD: University Park Press.

Fennell, C. T., Byers-Heinlein, K., & Werker, J. F. (2007). Using speech sounds to guide word learning: The case of bilingual infants. *Child Development, 78*(5) 1510-1525.

Hollich, G., Newman, R. S., & Jusczyk, P. W. (2005). Infants' use of synchronized visual information to separate streams of speech. *Child Development, 76*(3), 598-613.

Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, *44*(4), 548-567.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics, 50(2),* 93-107.

Kuhl, P. K. (1993). Innate predispositions and the effects of experience in speech perception: the

native language magnet theory. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P.

McNeilage, & J. Morton (Eds.) *Developmental neurocognition: Speech and face*

processing in the first year of life, NATO ASI Series (pp. 259-274). The Netherlands:

Kluwer Academic Publishers.

Munhall, K. G., Kroos, C., Jozan, G., & Vatikiotis-Bateson, E. (2004). Spatial frequency

requirements for audiovisual speech perception. *Perception & Psychophysics, 66*(4), 574-

583.

Navarra, J., & Soto-Faraco, S. (2007). Hearing lips in a second language: Visual articulatory

information enables the perception of L2 sounds. *Psychological Research, 71,* 4-12.

Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English-

learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and*

Language, *43*(1), 1-19.

Ramus, F. (2002). Language discrimination by newborns: Teasing apart phonotactic, rhythmic,

and intonational cues. *Annual Review of Language Acquisition, 2*, 85-115.

Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study

based on speech resynthesis. *Journal of the Acoustical Society of America, 105*(1), 512-

521.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal.

Cognition, 73(3), 265-292.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants.

Science, *274*, 1926-1928.

Soto-Faraco, S., Navarra, J., Weikum, W. M., Vouloumanos, A., Sebastián-Gallés, N., & Werker, J. F. (2007). Discriminating languages by speechreading. *Perception and Psychophysics, 69,* 218-231.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Statistical and stress cues in infant word segmentation. *Developmental Psychology, 39*, 706-716.

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

Appendix A[6]

The following represents a complete list of the sentences read from *The Little Prince/Le Petit Prince* used to create silent video clips of speakers reciting sentences in both French and English. The following sentences were spoken by three bilingual speakers and were shown to the adult participants in the studies discussed in Chapter III. Sentences were randomly chosen for each participant by the computer program designed for running the adult study.

English

1. For I do not want any one to read my book carelessly. I have suffered too much grief in setting down these memories.

2. Six years have already passed since my friend went away from me, with his sheep. If I try to describe him here it is to make sure that I shall not forget him.

3. To forget a friend is sad. Not every one has had a friend. And if I forget him, I may become like the grown-ups who are no longer interested in anything but figures...

4. It is for that purpose, again, that I have bought a box of paints and some pencils. It is hard to take up drawing again at my age.

5. I have never made any pictures except those of the boa constrictor from the outside and the boa constrictor from the inside, since I was six.

6. I shall certainly try to make my portraits as true to life as possible. But I am not at all sure of success.

7. One drawing goes along all right, and another has no resemblance to its subject. I make some errors, too, in the little prince's height: in one place he is too tall and in another too short.

---

[6] All sentences listed in appendix A were taken from the following book: de Saint-Exupéry, A. (1943). *Le Petit Prince/The Little Prince*. Translated by Katherine Woods. France: Gallimard.

8. And I feel some doubts about the colour of his costume. So I fumble along as best I can, now good, now bad, and I hope generally fair-to-middling.

9. In certain more important details I shall make mistakes, also. But that is something that will not be my fault. My friend never explained anything to me.

10. He thought, perhaps, that I was like himself. But I alas, do not know how to see sheep through the walls of boxes. Perhaps I am a little like the grown-ups. I have had to grow old.

11. I soon learned to know this flower better. On the little prince's planet the flowers had always been very simple. They had only one ring of petals.

12. One morning they would appear in the grass, and by night they would have faded peacefully away.

13. The little prince had watched very closely over this small sprout which was not like any other small sprout on this planet.

14. But the shrub soon stopped growing, and began to get ready to produce a flower. The little prince was present at the first appearance of a huge bud.

15. He felt at once that some sort of miraculous apparition must emerge from it. But the flower was not satisfied to complete the preparations of her beauty in the shelter of her green chamber.

16. She did not wish to go out into the world all rumpled, like the field poppies. It was only in the full radiance of her beauty that she wished to appear.

17. Oh, yes! She was a coquettish creature! And her mysterious adornment lasted for days and days. Then one morning, exactly at sunrise, she suddenly showed herself.

18. I believe that for his escape he took advantage of the migration of a flock of wild birds. On the morning of his departure he put his planet in perfect order.

19. He carefully cleaned out his active volcanoes. He possessed two active volcanoes; and they were very convenient for heating his breakfast in the morning.

20. He also had one volcano that was extinct. But, as he said, "One never knows!" so he cleaned out the extinct volcano, too.

21. If they are cleaned out, volcanoes burn slowly and steadily, without any eruptions. Volcanic eruptions are like fire in a chimney.

22. On our earth we are obviously much too small to clean out our volcanoes. That is why they bring no end of trouble upon us.

23. The little prince also pulled up, with a certain sense of dejection, the last little shoots of baobabs. He believed that he would never want to return.

24. But on this morning all these familiar tasks seemed very precious to him. And when he watered the flower for the last time, and prepared to place her under the shelter of her glass globe, he realized that he was very close to tears.

25. To give you an idea of the size of Earth, I will tell you that before the invention of electricity it was necessary to maintain over the whole of the six continents a veritable army of lamplighters for the street lamps.

26. Seen from a slight distance, that would make a splendid spectacle. The movements of this army would be regulated like those of the ballet in the opera.

27. First would come to turn of the lamplighters of New Zealand and Australia. Having set their lamps alight, these would go off to sleep.

28. Next, the lamplighters of China and Siberia would enter for their steps in the dance, and then they too would be waved back into the wings.

29. After that would come the turn of the lamplighters of Russia and the Indies; then those of Africa ad Europe; then those of South America; then those of North America.

30. And never would they make a mistake in the order of their entry upon the stage. It would be magnificent.

31. When one wishes to play the wit, he sometimes wanders a little from the truth. I have not been altogether honest in what I have told you about the lamplighters.

32. And I realize that I run the risk of giving a false idea of our planet to those who do not know it. Men occupy a very small place upon the Earth.

33. If the two billion inhabitants who people the surface were all to stand upright, all humanity could be piled up on a small Pacific islet.

34. The grown-ups, to be sure, will not believe you when you tell them that. They imagine that they fill a great deal of space. They fancy themselves as important as baobabs.

35. You should advise them, then, to make their own calculations. They adore figures, and they will please them. But do not waste your time on this extra task.

36. It is unnecessary. You have, I know, confidence in me. When the little prince arrived on the Earth, he was very much surprised not to see any people.

37. He was beginning to be afraid he had come to the wrong planet, when a coil of gold, the colour of the moonlight, flashed across the sand.

French

1. Car je n'aime pas qu'on lise mon livre à la 1égère. J'éprouve tant de chagrin à raconter ces souvenirs.

2. Il y a six ans déjà que mon ami s'en est allé avec son mouton. Si j'essaie ici de le décrire, c'est afin de ne pas 1'oublier.

3. C'est triste d'oublier un ami. Tout le monde n'a pas eu un ami. Et ie puis devenir comme les grandes personnes qui ne s'intéressent plus qu'aux chiffres.

4. C'est donc pour ça encore que j'ai acheté une boîte de couleurs et des crayons. C'est dur de se remettre au dessin, à mon âge.

5. On n'a jamais fait d'autres tentatives que celle d'un boa fermé et celle d'un boa ouvert, I' âge de six ans!

6. J'essaierai, bien sûr, de faire des portraits le plus ressemblant possible. Mais je ne suis pas tout à fait certain de réussir.

7. Un dessin va, et 1'autre ne ressemble plus. Je me trompe un peu aussi sur la taille. Ici le petit prince est trop grand. Là il est trop petit.

8. J'hésite aussi sur la couleur de son costume. Alors je tâtonne comme ci et comme ça, tant bien que mal.

9. Je me tromperai enfin sur certains détails plus importants. Mais ça, il faudra me le pardonner. Mon ami ne donnait jamais d'explications.

10. Il me croyait peut- être semblable à lui. Mais moi, malheureusement, je ne sais pas voir les moutons à travers les caisses. Je suis peut-être un peu comme les grandes personnes. J'ai dû vieillir.

11. J'appris bien vite à mieux connaître la fleur. Il y en avait toujours sur la planète du petit prince les fleurs très simples.

12. Elles étaient ornées d'un rang de pétales. Elles apparaissaient un matin dans l'herbe, et puis elles s'éteignaient le soir.

13. Le petit prince avait surveillé de très près cette brindille qui ne ressemblait pas aux autres brindilles.

14. Mais l'arbuste cessa vite de croître, et commença de préparer une fleur. Le petit prince assistait à l'installation d'un bouton énorme.

15. Il sentait bien qu'il en sortirait une apparition miraculeuse, mais la fleur n'en finissait pas de se préparer à étre belle, à l'abri de sa chambre verte.

16. Elle ne voulait pas sortir toute fripée comme les coquelicots. Elle ne voulait apparaître que dans le plein rayonnement de sa beauté.

17. Eh! Oui. Elle était très coquette! Sa toilette mystérieuse avait donc duré des jours et des jours. Et puis voici qu'un matin, justement à l'heure de lever du soleil, elle s'était montrée.

18. Je crois qu'il profita, pour son évasion, d'une migration d'oiseaux sauvages. Au matin du départ il mit sa planète bien en ordre.

19. Il ramona soigneusement ses volcans en activité. Il possédait deux volcans en activité. Et c'était bien commode pour faire chauffer le petit déjeuner du matin.

20. Il possédait aussi un volcan éteint. Mais, comme il disait : "On ne sait jamais! " Il ramona donc également le volcan éteint.

21. S'ils sont bien ramonés, les volcans brûlent doucement et régulièrement, sans éruptions. Les éruptions volcaniques sont comme des feux de cheminée.

22. Évidemment sur notre terre nous sommes beaucoup trop petits pour ramoner nos volcans. C'est pourquoi ils nous causent des tas d'ennuis.

23. Le petit prince arracha aussi, avec un peu de mé1ancolie, les dernières pousses de baobabs. Il croyait ne jamais devoir revenir.

24. Mais tous ces travaux familiers lui parurent, ce matin-là, extrêmement doux. Et, quand il arrosa une dernière fois la fleur, et se prépara à la mettre à 1'abri sous son globe, il se découvrit 1'envie de pleurer.

25. Pour vous donner une idée des dimensions de la Terre je vous dirai qu'avant 1'invention de 1'é1ectricité on y devait entretenir, sur 1'ensemble des six continents, une véritable armée d'allumeurs de réverbères.

26. Vu d'un peu loin ça faisait un effet splendide. Les mouvements de cette armée étaient réglés, comme ceux d'un ballet d'opéra.

27. D'abord venait le tour des allumeurs de réverbères de Nouvelle-Zé1ande et d'Australie. Puis ceux-ci, ayant allumé leurs lampions, s'en allaient dormir.

28. Alors entraient à leur tour dans la danse les allumeurs de réverbères de Chine et de Sibérie. Puis eux aussi s'escamotaient dans les coulisses.

29. Alors venait le tour des allumeurs de réverbères de Russie et des Indes. Puis de ceux d'Amérique du Sud. Puis de ceux d'Amérique du Nord.

30. Et jamais ils ne se trompaient dans leur ordre d'entrée en scène. C'était grandiose.

31. Quand on veut faire de 1'esprit, il arrive que 1'on mente un peu. Je n'ai pas été très honnête en vous parlant des allumeurs de réverbères.

32. Je risque de donner une fausse idée de notre planète à ceux qui ne la connaissent pas. Les hommes occupent très peu de place sur la terre.

33. Si les deux milliards d'habitants qui peuplent la terre se tenaient debout et un peu serrés, on pourrait entasser l'humanité sur le moindre petit îlot du Pacifique.

34. Les grandes personnes, bien sûr, ne vous croiront pas. Elles s'imaginent tenir beaucoup de place. Elles se voient importantes comme des baobabs.

35. Vous leur conseillerez donc de faire le calcul. Elles adorent les chiffres, ça leur plaira. Mais ne perdez pas votre temps à ce pensum.

36. C'est inutile. Vous avez confiance en moi. Le petit prince, une fois sur terre, fut donc bien surpris de ne voir personne.

37. Il avait déjà peur de s'être trompé de planète, quand un anneau couleur de lune remua dans le sable.

The following sentences represent the sentences chosen for the infant studies discussed in Chapter II and Chapter IV.

Speaker 1

English Sentences

-For I do not want any one to read my book carelessly. I have suffered too much grief in setting down these memories.

-It is for that purpose, again, that I have bought a box of paints and some pencils. It is hard to take up drawing again at my age.

-I have never made any pictures except those of the boa constrictor from the outside and the boa constrictor from the inside, since I was six.

-I shall certainly try to make my portraits as true to life as possible. But I am not at all sure of success.

-And I feel some doubts about the colour of his costume. So I fumble along as best I can, now good, now bad, and I hope generally fair-to-middling.

-One morning they would appear in the grass, and by night they would have faded peacefully away.

-The little prince had watched very closely over this small sprout which was not like any other small sprout on this planet.

-When one wishes to play the wit, he sometimes wanders a little from the truth. I have not been altogether honest in what I have told you about the lamplighters.

-It is unnecessary. You have, I know, confidence in me. When the little prince arrived on the Earth, he was very much surprised not to see any people.

Speaker 1

French Sentences

-J'essaierai, bien sûr, de faire des portraits le plus ressemblant possible. Mais je ne suis pas tout à fait certain de réussir.

-Elle ne voulait pas sortir toute fripée comme les coquelicots. Elle ne voulait apparaître que dans le plein rayonnement de sa beauté.

-Eh! Oui. Elle était très coquette! Sa toilette mystérieuse avait donc duré des jours et des jours. Et puis voici qu'un matin, justement à 1'heure de lever du soleil, elle s'était montrée.

-Il possédait aussi un volcan éteint. Mais, comme il disait : "On ne sait jamais! " Il ramona donc également le volcan éteint.

-S'ils sont bien ramonés, les volcans brûlent doucement et régulièrement, sans éruptions. Les éruptions volcaniques sont comme des feux de cheminée.

-Le petit prince arracha aussi, avec un peu de mé1ancolie, les dernières pousses de baobabs. Il croyait ne jamais devoir revenir.

-Quand on veut faire de 1'esprit, il arrive que 1'on mente un peu. Je n'ai pas été très honnête en vous parlant des allumeurs de réverbères.

-C'est inutile. Vous avez confiance en moi. Le petit prince, une fois sur terre, fut donc bien surpris de ne voir personne.

-Il avait déjà peur de s'être trompé de planète, quand un anneau couleur de lune remua dans le sable.

Speaker 2

English Sentences

-Six years have already passed since my friend went away from me, with his sheep. If I try to describe him here it is to make sure that I shall not forget him.

-In certain more important details I shall make mistakes, also. But that is something that will not be my fault. My friend never explained anything to me.

-I soon learned to know this flower better. On the little prince's planet the flowers had always been very simple. They had only one ring of petals.

-She did not wish to go out into the world all rumpled, like the field poppies. It was only in the full radiance of her beauty that she wished to appear.

-He carefully cleaned out his active volcanoes. He possessed two active volcanoes; and they were very convenient for heating his breakfast in the morning.

-On our earth we are obviously much too small to clean out our volcanoes. That is why they bring no end of trouble upon us.

-The little prince also pulled up, with a certain sense of dejection, the last little shoots of baobabs. He believed that he would never want to return.

-And I realize that I run the risk of giving a false idea of our planet to those who do not know it. Men occupy a very small place upon the Earth.

-He was beginning to be afraid he had come to the wrong planet, when a coil of gold, the colour of the moonlight, flashed across the sand.

Speaker 2

French Sentences

-On n'a jamais fait d'autres tentatives que celle d'un boa fermé et celle d'un boa ouvert, I' âge de six ans!

-Je me tromperai enfin sur certains détails plus importants. Mais ça, il faudra me le pardonner. Mon ami ne donnait jamais d'explications.

-Il me croyait peut- être semblable à lui. Mais moi, malheureusement, je ne sais pas voir les moutons à travers les caisses. Je suis peut-être un peu comme les grandes personnes. J'ai dû vieillir.

-J'appris bien vite à mieux connaître la fleur. Il y en avait toujours sur la planète du petit prince les fleurs très simples.

-Mais l'arbuste cessa vite de croître, et commença de préparer une fleur. Le petit prince assistait à l'installation d'un bouton énorme.

-D'abord venait le tour des allumeurs de réverbères de Nouvelle-Zé1ande et d'Australie. Puis ceux-ci, ayant allumé leurs lampions, s'en allaient dormir.

-Alors entraient à leur tour dans la danse les allumeurs de réverbères de Chine et de Sibérie. Puis eux aussi s'escamotaient dans les coulisses.

-Alors venait le tour des allumeurs de réverbères de Russie et des Indes. Puis de ceux d'Amérique du Sud. Puis de ceux d'Amérique du Nord.

-Je risque de donner une fausse idée de notre planète à ceux qui ne la connaissent pas. Les hommes occupent très peu de place sur la terre.

Speaker 3

English Sentences

-One drawing goes along all right, and another has no resemblance to its subject. I make some errors, too, in the little prince's height: in one place he is too tall and in another too short.

-He felt at once that some sort of miraculous apparition must emerge from it. But the flower was not satisfied to complete the preparations of her beauty in the shelter of her green chamber.

-Oh, yes! She was a coquettish creature! And her mysterious adornment lasted for days and days. Then one morning, exactly at sunrise, she suddenly showed herself.

-I believe that for his escape he took advantage of the migration of a flock of wild birds. On the morning of his departure he put his planet in perfect order.

-If they are cleaned out, volcanoes burn slowly and steadily, without any eruptions. Volcanic eruptions are like fire in a chimney.

-Seen from a slight distance, that would make a splendid spectacle. The movements of this army would be regulated like those of the ballet in the opera.

-Next, the lamplighters of China and Siberia would enter for their steps in the dance, and then they too would be waved back into the wings.

-If the two billion inhabitants who people the surface were all to stand upright, all humanity could be piled up on a small Pacific islet.

-The grown-ups, to be sure, will not believe you when you tell them that. They imagine that they fill a great deal of space. They fancy themselves as important as baobabs.

Speaker 3

French Sentences

-Il y a six ans déjà que mon ami s'en est allé avec son mouton. Si j'essaie ici de le décrire, c'est afin de ne pas 1'oublier.

-C'est donc pour ça encore que j'ai acheté une boîte de couleurs et des crayons. C'est dur de se remettre au dessin, à mon âge.

-Un dessin va, et 1'autre ne ressemble plus. Je me trompe un peu aussi sur la taille. Ici le petit prince est trop grand. Là il est trop petit.

-Il sentait bien qu'il en sortirait une apparition miraculeuse, mais la fleur n'en finissait pas de se préparer à étre belle, à 1'abri de sa chambre verte.

-Je crois qu'il profita, pour son évasion, d'une migration d'oiseaux sauvages. Au matin du départ il mit sa planète bien en ordre.

-Il ramona soigneusement ses volcans en activité. Il possédait deux volcans en activité. Et c'était bien commode pour faire chauffer le petit déjeuner du matin.

-Évidemment sur notre terre nous sommes beaucoup trop petits pour ramoner nos volcans. C'est pourquoi ils nous causent des tas d'ennuis.

-Vu d'un peu loin ça faisait un effet splendide. Les mouvements de cette armée étaient réglés, comme ceux d'un ballet d'opéra.

-Les grandes personnes, bien sûr, ne vous croiront pas. Elles s'imaginent tenir beaucoup de place. Elles se voient importantes comme des baobabs.

The University of British Columbia
Office of Research Services
Behavioural Research Ethics Board
Suite 102, 6190 Agronomy Road, Vancouver, B.C. V6T 1Z3

CERTIFICATE OF APPROVAL- MINIMAL RISK RENEWAL

PRINCIPAL INVESTIGATOR:	DEPARTMENT:	UBC BREB NUMBER:
Janet F. Werker	UBC/Arts/Psychology, Department of	H04-80441

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:	
Institution	Site
UBC Children's and Women's Health Centre of BC (incl. Sunny Hill) Other locations where the research will be conducted: N/A	Point Grey Site Children's and Women's Health Centre of BC (incl. Sunny Hill)

CO-INVESTIGATOR(S):
Laurel Fais
Whitney M. Weikum
Katherine Yoshida
Henny Yeung
Krista Byers-Heinlein
Vivian Yi Pan
Naz Akmal
Clarisa Markel

SPONSORING AGENCIES:
James S. McDonnell Foundation - "Program to Combine Methods for Assessing Cognitive Development in Human Infants and to Promote the Field of Developmental Cognitive Neuroscience" Social Sciences and Humanities Research Council of Canada (SSHRC) - "Bilingualism and Enculturation: Roots in Infancy"

PROJECT TITLE:
Bilingualism and Enculturation: Roots in Infancy

EXPIRY DATE OF THIS APPROVAL: June 7, 2008

APPROVAL DATE: June 7, 2007

The Annual Renewal for Study have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.

Approval is issued on behalf of the Behavioural Research Ethics Board