Acoustic-perceptual salience and developmental speech perception

by

Chandan Raghava Narayan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Linguistics)
in The University of Michigan
2006

Doctoral Committee:

    Professor Patrice Speeter Beddor, Co-Chair
    Professor Janet F. Werker, Co-Chair, University of British Columbia
    Professor Susan A. Gelman
    Professor Marilyn J. Shatz

For my mother and father

Table of Contents

List of Figures

List of Tables

CHAPTER 1: Introduction

1.1   Introduction

Speech communication involves a chain of physical and psychological events. The speaker initiates the chain by setting into motion the vocal apparatus, which propels a modulated airstream to the listener. The process culminates with the listener receiving the fluctuations in air pressure through the auditory mechanism and subsequently parsing the signal into higher-level linguistic units. This dissertation is concerned with the intersections of the final links in this chain, namely the resolution of speech acoustics by the listener, and the influence linguistic experience has on perception. I approach this moment in communicative interaction by examining the perceptual abilities of adults and infants whose linguistic backgrounds differ, therefore providing a window into those aspects of speech perception which are attributed to experience and those attributed to more general characteristics of the processing of speech acoustics.

In this dissertation I focus on how adults and infants perceive nasal consonants and whether their perceptual patterns are consistent with well-established patterns of perception of native and non-native sounds in adulthood and infancy in general. Why *nasal consonant* perception? The answer lies in the phonological typology of nasal

places of articulation in the world's languages, which show an asymmetry in that nearly all languages have /n/ and /m/ in syllable-onset position, while considerably fewer have /ŋ/. Prompted by the idea that the phonological distribution of nasals appears as it does, in part, because some nasal place contrasts are less salient than others, I investigate whether this salience affects the perceptual development of these contrasts in human listeners. In particular, I ask whether perceptually less salient contrasts might show a pattern of developmental speech perception different from the well-known tendency for phonetic contrasts to be discriminable in young infancy and to reorganize during the first year so the non-native contrasts are no longer discriminable at 10-12 months (Werker & Tees, 1984a). I conduct this investigation with an eye towards two—not incompatible—working hypotheses concerning how infants discriminate native and non-native nasal contrasts. The first hypothesis is that infants' perception of non-native nasal contrasts follows a *reorganization* pattern. The second hypothesis is that infants are sensitive to the *acoustic-perceptual salience* of nasal consonant contrasts, and this salience affects infants' discrimination of these contrasts.

The remainder of this introductory chapter seeks to motivate the two hypotheses by reviewing relevant literature supporting a reorganization hypothesis for infants' perception of non-native nasal consonant contrasts and evidence from perceptual explanations for typological patterns suggestive of an acoustic-perceptual salience account. I conclude the introduction by offering an outline of the subsequent chapters of the dissertation.

1.2   The reorganization hypothesis

The *reorganization* hypothesis follows from decades of research into the speech perception abilities of adults and infants in general, and their perception of non-native speech contrasts in particular. The central motivation for much of this research is the well-established finding that adult speech perception is affected by native-language phonology, in that non-native speech contrasts are often difficult for adults to discriminate. The reorganization hypothesis suggests that, when adults show difficulties discriminating a non-native speech contrast, infants' discrimination of the same contrast will develop from being good in young infancy (or language-general perception) into being adult-like (marked by less successful discrimination) in later infancy, thus reflecting the contrasts employed by their ambient language. Support for this hypothesis is discussed below.

1.2.1   Adults show native-language effects in speech perception

The extensive literature on adult non-native speech perception, which has been reviewed by Strange (1995), Harnsberger (1998, 2000), and Cho and McQueen (2006) among others, converges on the finding that adults are influenced by their native-language phonology in their perception of non-native speech contrasts. This general finding is characterized by two classic studies from the 1970s. Abramson and Lisker (1970) showed that listeners' identification and discrimination of stimuli from a voice onset time (VOT) continuum coincided with the boundaries shown in the acoustic analyses of their respective languages. Both Spanish and English speakers' responses showed a single discrimination peak corresponding to the voiced-voiceless distinction, which is phonological in both languages; Thai speakers' responses, on the other hand,

3

showed two peaks corresponding to the distinctions between voiced, voiceless unaspirated and voiceless aspirated categories, which are phonemic in Thai. This cross-language paradigm was extended to other phonetic features, such as those distinguishing the American English liquids /ɹ/ and /l/. In a second, now-classic study, Miyawaki *et al.* (1975) presented American English and Japanese speakers with a synthetic [ɹ]-[l] continuum in which only the F3 transitions varied. American English listeners accurately discriminated continuum members that crossed their [ɹ]-[l] identification boundary and poorly discriminated within categories. Consistent with Japanese phonology, in which the liquids /ɹ/ and /l/ do not contrast but /ɹ/ is realized postvocalically as a flap, Japanese listeners showed poor performance on all discrimination pairs.

1.2.2   Infant perception of non-native speech contrasts

The results of adult cross-language speech perception studies raised the question of the nature of the development of speech perception of infants. The infant research in the 1970s provided an interesting complement to the adult findings. While adults show the influence of their native language in their discrimination of non-native contrasts, research began to accumulate suggesting that infants are able to discriminate *both* native and non-native contrasts equally well. Following classics studies by Eimas *et al.* (1971), which showed that American English-hearing infants as young as 1-2 months are able to discriminate between-category stimuli along a VOT continuum, and Kuhl (1979), which showed that infants can generalize across talker and pitch to discriminate vowel categories, Lasky *et al.* (1975), Streeter (1976) and Trehub (1976) provided evidence that infants are capable of discriminating non-native speech contrasts regardless of the

4

phonological structure of their ambient language. In general, these studies suggested that the native-language effects seen in adult speech perception seem to not affect infant speech perception.

With the experimental data suggesting that infants perceive many speech contrasts in a *language-general* manner, while adults are influenced by their native-language phonology, Werker and her colleagues began a program to chart the time course over which humans begin to perceive speech like adults. In a developmental study comparing 6-8 month old English-hearing infants to both English-speaking and Hindi-speaking adults, Werker *et al.* (1981) confirmed the general finding that young infants can discriminate a non-native consonant contrast which adults have difficulty discriminating. Using a conditioned-head-turn (CHT) procedure, Werker and colleagues found that English-hearing infants at 6-8 months and Hindi-speaking adults were able to discriminate two Hindi contrasts, voiceless dental vs. retroflex stops ([t̪a]-[ʈa]) and breathy voiced vs. voiceless aspirated ([d̪ʰa]-[t̪ʰa]) stops. Consistent with the general finding in adult non-native speech perception, English-speaking adults did not reliably discriminate the contrasts.

In a follow up study, Werker and Tees (1983) found that, like English-speaking adults, older English-speaking children aged 4-, 8-, and 12-years were all poor at discriminating the Hindi contrasts. Werker and Tees (1984a) then looked at three infant age groups (6-8, 8-10, and 10-12 months) and their perception of two non-native contrasts, the Hindi dental-retroflex contrast and a velar-uvular place contrast in Nthlakampx (Salish) ejectives ([k'i]-[q'i]). Using the CHT task, they found that the 6-8 month-old English infants successfully discriminated the non-native contrasts, while the

10-12 month-old infants did not. Additionally, 10-12 month-old infants from Hindi-speaking and Nthlakampx-speaking households successfully discriminated their respective *native* contrasts. This study not only located the 10-12 month range as the age when infants perceive non-native consonant contrasts like adults, but also affirmed that the change is due to linguistic experience accrued over the course of the first year of life. Werker and Tees (1984a) termed this development, from language-general abilities to discrimination tuned to the structure of the ambient language, a perceptual "reorganization." Evidence for a reorganization of infants' perceptual space has been found subsequently for vowel contrasts (Polka & Werker, 1994), as well as other consonant contrasts (Best, 1993; Tsushima *et al.*, 1994; Anderson *et al.*, 2003).

Interestingly, other experimental findings suggested that not all non-native consonant contrasts follow the reorganization pattern. Best *et al.* (1988) presented 12-14 month-old English infants and English-speaking adults with Isizulu (Bantu) click contrasts. In a visual habituation task *both* adults and infants successfully discriminated the click contrasts. Best and colleagues reasoned that the click contrasts were sufficiently different from English contrasts (being perceived as nonspeech) to garner accurate discrimination, even when the infant's language experience should restrict her in discriminating only those contrasts which are functional in the ambient speech.

In general, however, there is overwhelming evidence suggesting that non-native speech perception becomes reorganized from infancy to adulthood. When adults show effects of non-native speech perception (i.e., poor discrimination of non-native speech contrasts), infants show a pattern of development whereby initially discriminable contrasts are not reliably discriminated by the end of their first year. If *nasal* consonant

perception proceeds in a manner similar to the perception of other consonants described in the infant speech perception literature, we might expect a reorganization in infants' perception of non-native nasal place contrasts. The reorganization hypothesis predicts that infants will reliably discriminate non-native nasal contrasts at 6-8 months, but succeed at discriminating only native nasal contrasts at 10-12 months.

The pattern of reorganization becomes complicated when we consider the relative acoustic salience of speech contrasts. In the next section I suggest that not all non-native consonant contrasts should be considered equal. Rather, contrasts that are typologically rare potentially arise from weak acoustic-perceptual salience and this may affect the development of speech perception.

1.3 The acoustic-perceptual salience hypothesis

The *acoustic-perceptual salience* hypothesis predicts that speech contrasts that are particularly difficult to differentiate will have consequences for developmental speech perception. Perceptual difficulty is assessed in this study by presenting adults with two native nasal place contrasts. If adults are better at discriminating one nasal place contrast than another, then the easier contrast is more "salient" than the contrast that posed more difficulty for the listeners. If significant effects of acoustic-perceptual salience are found in adult perception, then I argue that corresponding effects will be observed in infant perception as well. Under the acoustic-perceptual salience hypothesis, infants are predicted to discriminate the less salient contrast less accurately than the more salient contrast. I suggest in the following section that a potential consequence of relative acoustic-perceptual salience of nasals can be found in the distribution and frequency of nasal places of articulation in the world's phonologies. I present a case for an acoustic-

7

perceptual source to the place of articulation asymmetry in the distribution of nasal consonants in the world's languages and suggest that nasal place perception provides an appealing test case for the investigation of acoustic-perceptual salience effects in infant speech perception.

1.3.1   Nasal place typology

Ferguson's (1963) basic typology of nasals makes generalizations regarding the distribution of *Primary Nasal Consonants* (PNC)[1] in the world's languages. He claims that:

1. If in a given language there is only one PNC, it is /n/ (or the apical nasal)

2. If in a given language there are only two PNCs, they are /m/ (bilabial) and /n/

The asymmetry that emerges in nasal consonant distribution in the world's languages is that phonologies tend to have /n/ and /m/ over post-alveolar places of articulation.

The asymmetry is empirically confirmed in Maddieson's (1984) classic survey of the sound patterns of 317 of the world's languages in the UCLA Phonological Segment Inventory Database (UPSID). In that database, 316 languages have apical nasals (at dental, dental/alveolar, or alveolar places of articulation), 299 have the bilabial nasal and 167 have the velar. Maddieson noted that the presence of an oral stop at a particular place of articulation does not necessarily imply a homorganic nasal. He found that while 283 languages in the database exhibited the voiceless oral velar obstruent /k/, only 167 showed the homorganic nasal stop /ŋ/ (Maddieson, 1984).

---

[1] Ferguson defines PNC as "a phoneme of which the most characteristic allophone is a voiced nasal stop, that is, a sound produced by a complete oral stoppage (e.g., apical, labial), velic opening, and vibration of the vocal cords" (Ferguson, 1963).

The typology becomes more complex when we consider the intra-syllable distribution of the velar nasal. In a recent survey of 468 genetically diverse languages (Haspelmath *et al.*, 2005), Anderson (2005) found that half lacked the velar nasal altogether (in both syllable-onset and coda positions). If a language had the velar nasal, it most often lacked it in syllable-onset position: of the 234 languages exhibiting the velar nasal, 88 had it in syllable-coda position only, while the remaining 146 had it in *both* onsets and codas. There was only one language in Anderson's survey, Grebo (Kru, Niger-Congo; Liberia), in which /ŋ/ occurs in *only* syllable-onset position. Taken together, the positional typology suggests that if a language has the velar nasal, but lacks it in one syllabic position, it most likely does so in syllable onsets.

1.3.2   An acoustic-perceptual source for phonological patterns

The typology of nasal place clearly exhibits an asymmetry. The next logical question to ask is *why* /n/ and /m/ are found more often than /ŋ/ in the phonologies of the world. The homorganic oral stops /p t k/ and /b d g/ do not show a similarly strong asymmetry. A large body of literature within the broader field of linguistic phonetics has sought to address the role played by phonetics in the shaping of recurring patterns in the phonologies of the world's languages. Some phonological patterns are said to have a source in phonetic properties arising from physiological constraints on vocal production while others are linked to the ways humans perceive the resulting acoustics. Ohala (1981 and elsewhere) directly attributes patterns in sound change to the perceiver. His proposal that the listener is the source of sound change follows from the notion that phonological patterns be deductively supported based on the operating principles of how speech is produced and the potential for the listener to misperceive the speaker's intended

9

production. In fact, there is substantial experimental evidence that phonological

patterns—including phonological patterns involving nasals—surface, in part, due to

perceptual factors (e.g., Kawasaki, 1986; Beddor *et al.*, 2001). Numerous other

researchers have recognized and empirically measured the delicate interplay between

production, acoustics and perception as contributing to the overall shape of phonological

systems (e.g., Stevens, 1989; Lindblom, 1990). Given the large body of evidence

pointing towards perceptual explanations for phonological patterns, I approach the nasal

place asymmetry by suggesting that its source may be found in the relationship between

the acoustics of nasal place and its corresponding perception.

1.3.3   Potential effects of acoustic-perceptual salience on infant speech perception

It is in the spirit of perceptual explanations for phonological patterns that I couch

the present appeal to the acoustic-perceptual salience of nasal consonant contrasts as

potentially having import in the development of their perception. Recent work by Polka

et al. (2001) also hints at such a relationship. They examined the perception of the

English syllable-initial [d]-[ð] contrast by Canadian French- (CF) and Canadian English

(CE)-speaking adults and infants (at 6-8 and 10-12 months) growing up in French- or

English-speaking households; CF does not exhibit this contrast. They found that

although the adult listeners showed the classic pattern of non-native speech perception,

with CF adults poorly discriminating the English contrast, CF and CE infants in both age

groups performed equally well, with roughly 70% of infants in both groups reaching the

discrimination criterion in the CHT task. That is, ambient language did not affect

discrimination of this contrast during the first year, despite perceptual differences for the

adult language groups. Intersubject variability in CF infants' sensitivity (as measured by

A' scores) increased with age, while decreasing in CE infants. Polka *et al.* interpreted these results as suggesting a pattern of development whereby native-language experience facilitates perception, as CE infants' perception improved over the course of development, and CF infants' discrimination abilities diminished by 10-12 months. The source of this new pattern of development in speech perception was not directly investigated by Polka and her colleagues; rather, the authors speculated that infants' poor discrimination of the native contrast might be attributable to acoustic factors ([ð] especially being acoustically "weak") and distributional facts regarding /ð/ in English. Polka *et al.*'s appeal to the relatively low acoustic salience of the [d]-[ð] contrast is bolstered by the fact that previous studies have suggested that infants have difficulty discriminating contrasts involving non-sibilant fricatives (Eilers & Minifie, 1975; Eilers *et al.*, 1977).

Polka *et al.*'s (2001) findings provide an interesting accompaniment to the acoustic-perceptual salience hypothesis for they suggest that contrasts that are acoustically weak (in Polka et al.'s terms) may be difficult to resolve in infancy and only later become facilitated with native-language experience. Polka *et al.*'s findings lie in contrast to the theoretical framework proposed by Burnham (1986), which predicts that phonetic contrasts that are "fragile" (defined as low in psychoacoustic salience and rare among the world's languages) will become reorganized in infancy (from good discrimination in early infancy and less accurate discrimination in late infancy) if the ambient language does not contrast them. For example, Burnham suggests that the Hindi dental-retrofelx contrast from Werker and Tees (1984a) is discriminable to young English-hearing infants and no longer discriminated by 10-12 month olds precisely

because of the rarity and low acoustic salience of the contrast. Robust contrasts, on the other hand, which are defined as highly acoustically salient and common in the world's languages, are predicted to show good discrimination throughout infancy and into childhood even without specific-language experience. Although Burnham's (1986) framework does not accommodate Polka *et al.*'s (2001) finding, the fragile-robust delineation provides an important heuristic for the present investigation for it identifies the importance of acoustic-perceptual salience and typology in making predictions for infant discrimination.

If there is an acoustic-perceptual source to the phonological patterning of nasal place contrasts, we might speculate that contrasts that are rare in the world's languages, like the alveolar-velar nasal contrast in syllable-onset position described above, are less typical because of the difficulty involved in their perception. The perception of nasal contrasts involving a velar place distinction may show effects similar to those observed in the Polka *et al.* (2001) study.

1.4   The role of experience in infant speech perception

Although the reorganization and acoustic-perceptual salience hypotheses are set up as competing—albeit not mutually exclusive—theories making different predictions about infants' discrimination patterns, Aslin and Pisoni (1980) present a developmental framework that allows us to bridge the hypotheses. Following Gottlieb (1976), Aslin and Pisoni identify the major roles that early postnatal language experience can play in modifying the relative discriminability of speech sounds (Figure 1.1).
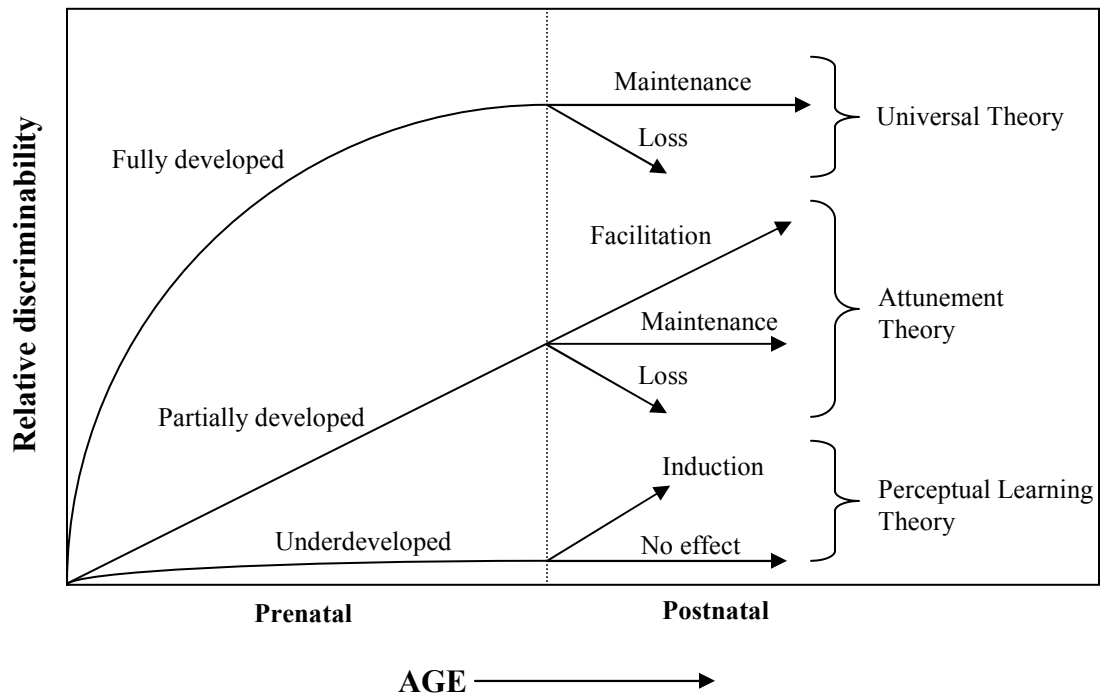
Figure 1.1 The major roles played by early postnatal language experience in speech perception. Adapted from Aslin and Pisoni (1980).

Aslin and Pisoni (1980) propose that infant speech perception has three potential starting points at the onset of (postnatal) linguistic experience: underdeveloped, partially developed, and fully developed. Underdeveloped perception suggests that the phonetic contrasts are poorly discriminated in neonates. A fully developed perception is one where phonetic contrasts are accurately discriminated at the onset of experience (i.e., without the influence of ambient language experience). Partially developed perception allows for a *middle ground* of perceptual ability in infancy—suggesting that some phonetic contrasts are well discriminated, while others are not. Aslin and Pisoni (1980) argue that infant speech perception can follow different routes to mature perception guided by phonetic and psychoacoustic factors. Each of these initial biases are affected

by experience in different ways. Under attunement theory, partially developed discrimination abilities are facilitated, maintained, or lost. Under the universal theory, fully developed abilities are either maintained or lost. The perceptual learning theory accounts for initially underdeveloped abilities and suggests that experience either induces discrimination or has no effect.

These different routes are not competing theories but rather characterize the role that experience with the native language plays in speech perception. By far the most well-attested theory accounting for patterns of developmental speech perception is attunement theory. Attunement posits that infants are initially capable of discriminating some phonetic contrasts but their abilities are broadly specified. Experience with the native language serves to sharpen initial perceptual boundaries, with relevant contrasts becoming more finely tuned (facilitation). Irrelevant contrasts remain broadly tuned (maintenance) or become attenuated (loss). Examples of developmental "loss" under the attunement theory include English-infants' perception of the Hindi dental/retroflex contrasts (Werker and Tees, 1984a) and German high rounded/unrounded vowels (Polka and Werker, 1994), and Japanese infants' perception of English liquids (Tsushima *et al.*, 1994). In general, there is overwhelming evidence from the infant speech perception literature for an attunement theory of developmental speech perception. Although Werker characterizes this pattern of development as "reorganization" rather than loss, the loss/reorganization pattern is the most common pattern in the infant cross-language speech perception literature. Polka *et al.*'s (2001) results (reviewed above) suggest facilitation of [d]-[ð] perception after the first year of development in Canadian-English hearing infants. More recently, Kuhl *et al.* (2006) report facilitation for the [ɹ]-[l] in the

second half of English-hearing infants' first year. In that study, English- and Japanese-hearing infants at 6-8 months discriminate the [ɹ]-[l] contrast equally accurately. By 10-12 months, however, English-hearing infants' performance accuracy increases significantly (consistent with English phonology), while Japanese-hearing infants' perception decreases (consistent with Japanese phonology).

The perceptual learning theory of speech perception development suggests that the ability to discriminate any particular contrast is specific to early experience with the ambient language. Under this theory the rate of development is dependent upon the frequency of the contrast, the psychophysical discriminability of the contrast and the infant's attentional state. Phonologically irrelevant contrasts are never discriminated *better* than native contrasts. There are no definitively identifiable cases of induction in the infant speech perception literature. Initially, Eilers *et al.*'s (1977) finding that infants have difficulty discriminating between certain fricative contrasts was taken as evidence for induction. But a follow-up study by Jusczyk *et al.*(1979) provided contrary evidence showing that infants discriminate fricatives ([fa]-[θa]) categorically but with the category boundary located at the adult [ba]-[da] boundary. This left Aslin and Pisoni (1980) to conclude that infants' discrimination of fricatives is initially based on the psychophysical (or acoustic salience) attributes of the signal, becoming refined over time, thus attunement may also account for the data.

Aslin and Pisoni's (1980) framework allows for a bridging of the acoustic-salience (psychophysical) and reorganization (phonetic) hypotheses in explaining speech perception development under the attunement theory. The acoustic-perceptual salience hypothesis, the effects of which are seen in the "facilitation" of [d]-[ð] in English

15

speakers between infancy and adulthood (Polka, *et al.*, 2001), and the reorganization

hypothesis, the effects of which are seen in the "maintenance" of the Hindi dental-

retroflex contrast in Hindi-hearing infants and its "loss" in English-hearing infants, are

both subsumed under attunement theory. Though Aslin and Pisoni do not directly

implicate acoustic-perceptual salience and reorganization as leading to divergent routes

within attunement theory, the effects of these two hypotheses, as seen in Polka *et al.*

(2001) and Werker and Tees (1984a), provide evidence that both hypotheses can be

unified under the rubric of attunement theory. In the concluding chapter I situate the

results of the experiments in the dissertation within this developmental framework.

1.5   Target language: Filipino

The choice of a target language for an investigation assessing the reorganization

and acoustic-perceptual salience hypotheses is crucially important. In order to test the

hypotheses put forth in 1.2 and 1.3, a target language exhibiting a typologically rare and

a typologically ubiquitous contrast needs to be chosen. The cross-language design allows

for an assessment of relative salience of contrasts within the target language as well as an

assessment of non-native effects. The next section introduces Filipino as the target

language for the adult and infant experiments. After a brief look at the phonology of

Filipino, I ask whether the general nasal place asymmetry exhibited across the world's

languages has a corresponding intra-language effect. That is, does Filipino, a language

that allows velar nasals in syllable onsets, nonetheless exhibit an asymmetry in their

distribution?

16

1.5.1   Phonological sketch of Filipino

Filipino[2] or Pilipino is a standardized language based on Tagalog (Austronesian, Malayo-Polynesian, Western Malayo-Polynesian, Meso Philippine, Central Philippine, Tagalog) and is the national language of the Republic of the Philippines. Filipino has roughly 22 million speakers in the Philippines alone, according to the 2000 Philippine census conducted by the National Statistics Office of the Philippines. Filipino is also spoken in the United States, Canada, Guam, Saudi Arabia, and the United Arab Emirates, among other countries. Speakers whose first language is Filipino/Tagalog reside mainly on the island of Luzon (including Metro Manila and Quezon) as well as on the islands of Lubang, Marinduque and Mindoro.

The phoneme inventory of Filipino is given in Table 1.1. Filipino has 16 consonant phonemes and 5 vowel phonemes (Carrier, 1979). Filipino exhibits the velar nasal phoneme in both syllable-onset and coda positions. The inventory is identical to that found in linguistic descriptions of Tagalog and similar to that of other Austronesian languages of the Philippines (Rubino, 2000).

Table 1.1. Consonant and vowel phoneme inventory for Filipino.

	Bilabial	Alveolar	Velar	Glottal
Plosive	**p** **b**	**t** **d**	**k** **g**	ʔ
Fricative		**s**		**h**
Nasal	**m**	**n**	**ŋ**	
Liquid		**l**		
Tap		**ɾ**		
Glide	**w**	**j**		

---

[2] The term "Filipino" is used rather than "Tagalog" in this dissertation as the latter term is also used to describe the Tagalog ethnic group. "Filipino" is considered a neutral designation as many of the speakers and listeners who participated in the experiments in this dissertation belong to other, non-Tagalog, ethnic communities.

	Front	Mid	Back
High	**i**		**u**
Mid	**e**		**o**
Low		**a**	

Filipino is rich in lexical borrowings from Spanish, Hokkien, Sanskrit and Arabic. As a result, sounds such as [ʃ], [tʃ], [dʒ], [ɚ], and [f] are found in loanwords (Zuraw, 2000). The typical syllabic structure in Filipino is CV(C), although onset and coda consonant clusters are sometimes found in loanwords. Stress is contrastive in Filipino, though there is some controversy as to whether the contrastive property is actually stress or vowel length (Zuraw, 2000).

1.5.2 Intra-language typology of Filipino nasals

In order to determine the frequency of the three Filipino syllable-onset nasals, /m/, /n/ and /ŋ/, in spontaneous speech, I examined the Filipino spontaneous speech corpus developed by the Digital Signal Processing Laboratory at the University of the Philippines, Diliman. The corpus contains transcriptions of recordings of the spontaneous speech of 60 native speakers of Filipino as transcribed by an Automatic Speech Recognition (ASR) system developed at the Laboratory. The corpus consists of speakers' story telling and their candid answers to a questionnaire.

I examined the corpus using a specially designed Perl script that searched transcribed text for instances of syllable-initial nasal consonants (marked in the corpus text as followed by a vowel) across varying vowel contexts. The resulting distribution is plotted in Figure 1.2.

18

Frequency of /NV/ syllables within a vocalic context
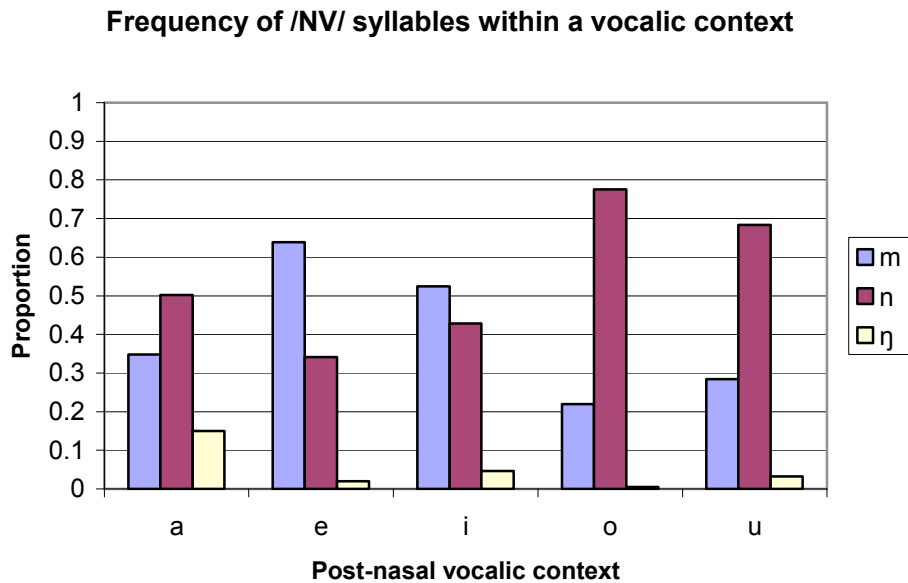
Figure 1.2 Frequency of syllable-onset nasal consonants in the spontaneous speech of 60 Filipino speakers.

The figure shows that before central (/a/) and back (/o/ and /u/) vowels the most frequent nasal is /n/. Before the front vowels (/i/ and /e/) the most frequent nasal is /m/. Across all vowel contexts, the least frequent nasal is /ŋ/[3]. Overall, the most frequently occurring syllable-onset nasal is alveolar, followed by the bilabial then the velar. That /ŋ/ constitutes only 6% of onset nasals in Filipino, while /n/ constitutes 54% and /m/ 39%, indicates a clear intra-language asymmetry in nasal place distribution.

In summary, the syllable-onset velar nasal in Filipino appears less frequently than either dental/alveolar or bilabial nasals in Filipino. This asymmetry resembles the cross-linguistic typology of nasal place, where the most common nasals in the world's languages are dental/alveolar followed by bilabial and then velar. Given that

---

[3] I conducted an analysis of a spontaneous speech corpus of Thai (LOTUS corpus from National Electronics and Computer Technology Center), which also has syllable-initial /ŋ/. Thai showed a similar asymmetry, with the velar-nasal onset occurring least frequently across vocalic contexts.

phonological patterns have been shown to reflect acoustic-perceptual relationships, Filipino is an appropriate target language for an inquiry into the role played by acoustic-perceptual salience in the development of nasal consonant perception. The potential effects of frequency on infants' formation of nasal categories is discussed in Chapter 7.

1.6   Summary and experimental design

To summarize the preceding sections, I approach adult and infant discrimination of nasal place contrasts (/n/, /m/ and /ŋ/) by focussing on two hypotheses for predicting and explaining discrimination patterns. The reorganization hypothesis follows from the findings of Werker and Tees (1984) and numerous other researchers who observed a perceptual reorganization in infants' ability to discriminate non-native contrasts towards the end of the first year of life. The reorganization occurs *only* when adults show an effect of native language in non-native contrast discrimination. The acoustic-perceptual salience hypothesis follows from research suggesting that phonological patterns appear the way they do in part due to acoustic and resulting perceptual factors. When a phonological contrast is relatively uncommon in the world's languages, we might reason that there is an articulatory or acoustic-perceptual source for its rarity. The acoustic-perceptual salience hypothesis suggests that the salience of speech contrasts, as measured in adult discrimination, potentially influences infant discrimination. A hint of this is seen in the findings of Polka *et al.* (2001) where discrimination of the English [d]-[ð] contrast was shown to be discriminated equally well by English- and French-hearing infants through the first year of life and better discriminated by older English-hearing infants.

The investigation begins with an assessment of the acoustic facts related to the nasal place asymmetry described in sections 1.3.1 and 1.3.2. Given the tendency for

20

phonologies of the world's languages to exhibit /n/ and /m/ more often than /ŋ/ in syllable-onset position, I ask whether there is an acoustic-perceptual source for the asymmetry by examining the acoustics the three nasals and their corresponding perception by adults. The acoustic-perceptual salience hypothesis is justified *only* if adults show an effect whereby an [n]-[m] contrast is more discriminable than [n]-[ŋ] contrast even when both contrasts are native. Such a result provides motivation for an investigation into infants' perception of nasal place contrasts under the rubric of the acoustic-perceptual salience hypothesis and the reorganization hypothesis. In examining infants' perception of nasal place contrasts I present native and non-native contrasts to three age groups (10-12, 6-8 and 4-5 months old) of infants, assessing infants' discrimination at each step based on the predictions made by the two working hypotheses. Previewing the results, I conclude that the acoustic-perceptual salience hypothesis better accounts for the discrimination pattern for native and non-native nasal place contrasts in infancy.

1.7   Outline of chapters

The dissertation proceeds as follows. I begin with an acoustic analysis (Experiment 1, Chapter 2) of syllable-onset nasals in Filipino in order to determine whether the asymmetric distribution of the different places of articulation has a corresponding asymmetry in terms of magnitude of acoustic differences. These analyses show that in the perceptually relevant F2xF3 space, [na] tokens are closer to [ŋa] than they are to [ma]. I next examine the perceptual reflex of this acoustic asymmetry as Experiment 2 in **Chapter 3**. In a cross-language design, I ask both English- and Filipino-speaking listeners to discriminate the pairs [ma]-[na] and [na]-[ŋa]. The results show that

English listeners successfully discriminate the native [ma]-[na] contrast, but perform at chance levels on [na]-[ŋa]. Filipino listeners discriminate both contrasts at rates above 90% correct, but perform significantly worse on [na]-[ŋa] than [ma]-[na], hinting at an acoustic-perceptual salience effect. These results motivate the next set of experiments, where I explore infants' perception of the perceptually salient [ma]-[na] contrast and the less salient [na]-[ŋa] contrast.

In **Chapter 4** I present English hearing infants at 10-12 (Experiment 3), 6-8 (Experiment 4), and 4-5 months (Experiment 5) with the native and non-native contrasts in a visual habituation task. The results of Experiments 2 and 3 show that English-hearing infants do not reliably discriminate the non-native [na]-[ŋa] contrast at either 10-12 or 6-8 months, but successfully discriminate the native [ma]-[na] contrast at both ages. That 6-8 month old infants were unable to discriminate the non-native contrast raised the question of whether even younger English-hearing infants might make the non-native discrimination. I then present 4-5 month old English-hearing infants with both the native [ma]-[na] and non-native [na]-[ŋa] contrasts (Experiment 5). As in Experiments 3 and 4, the 4-5 month old infants discriminated the native but not the non-native contrast. The failure of English hearing infants to discriminate the non-native [na]-[ŋa] contrast (even at ages when they are predicted to make the discrimination) suggests that the contrast is inherently different from the [ma]-[na] contrast.

In **Chapter 5**, I present Filipino-hearing infants at 10-12 and 6-8 months with the *native* [na]-[ŋa] contrast (Experiment 5). The Filipino group, who are expected to discriminate the contrast at both ages (based on the reorganization hypothesis), successfully make the discrimination only at 10-12 months. The results of the cross-

language infant experiments reinforce the notion that the [na]-[ŋa] contrast is inherently different from the [ma]-[na] contrast and that the overall pattern of development reflects this distinction. Experiment 6 in **Chapter 6** re-visits the argument that the [na]-[ŋa] contrast is perceptually less salient than the [ma]-[na] contrast. In this final experiment, I present adult Filipino listeners with both contrasts under degraded listening conditions. The results show that indeed Filipino listeners, for whom the [na]-[ŋa] contrast is linguistically salient, show an asymmetrical pattern of response when given the two contrasts in noisy conditions, with noise affecting the perception of [na]-[ŋa] more than the perception of [ma]-[na]. I conclude the dissertation in **Chapter 7** by discussing the results of the adult and infant experiments in light of the reorganization and acoustic-perceptual salience hypotheses. Finally, I propose a "next step" addressing how infants might learn phonetic contrasts (like the distinction between /n/ and /ŋ/) which are initially difficult to resolve.

CHAPTER 2: Acoustic characteristics of place of articulation in Filipino nasals

2.1   Introduction

In this chapter I examine the acoustic characteristics distinguishing place of articulation in three syllable-onset nasal consonants from Filipino, /m/, /n/ and /ŋ/. The purpose of the analysis is to determine the systematic acoustic differences among the three places of articulation across the productions of three speakers of Filipino. The results of this analysis will inform the interpretation of subsequent adult and infant perceptual experiments in this dissertation.

2.1.1   Acoustics of nasal place of articulation in syllable-initial position

The acoustic characteristics of nasal place have received considerable attention in the phonetics literature of the past 50 years (House, 1957; Fujimura, 1962; Kurowski & Blumstein, 1987; Seitz *et al.*, 1990; Qi & Fox, 1992). The majority of these studies have considered place-specific characteristics of two regions of the acoustic signal for nasals—the nasal murmur and the formant transition at the juncture with an adjacent vowel. Nasals are sonorants and necessarily involve both oral and nasal cavities as resonators. The spectral structure of the nasal is made up of resonances from the pharyngeal and

nasal cavities as well as anitiresonances, or zeros, from the oral cavity as predicted by the acoustic theory of speech production (Fant, 1960; Stevens, 1998). The juncture between the nasal consonant and surrounding vowels is similar to the formant transitional structure of their homorganic oral counterparts. The acoustic features of the nasal consonant are described in detail below.

The nasal *murmur* is made up of a low-frequency, relatively high amplitude resonance, sometimes called the nasal formant, and a series of higher frequency, low-amplitude formants. As a result of oro-nasal coupling, the spectrum of the murmur also consists of *antiformants* which essentially dampen the spectral energy around it. Following from acoustic theory, the place of articulation of the nasal stop should be predictable based on the frequency of the antiformant, which is determined by the dimensions of the side-branching resonator. In practice, however, prediction of nasal place based solely on the spectral characteristics of the murmur is difficult. Fujimura's (1962) study of nasal murmurs ([n], [m] and [ŋ] in /VNV/[4] contexts across various vowels) as produced by three male talkers suggested extensive intra- and inter-speaker variability in production. Variability also occurred within the murmur itself due to movement of the pole-zero pairs throughout the duration of the murmur. Nonetheless, he found a single antiformant at around 800 Hz for [m], around 1500-2000 Hz for [n] and above 3000 Hz for [ŋ]. Qi and Fox (1991) were also interested in determining those aspects of the nasal murmur which are directly related to place of articulation. In their study of NV syllables ([n] and[m] with varying vowel contexts) they found that by using

---

[4] Here and throughout the rest of the dissertation N=nasal consonant and V=vowel.

a perceptually based linear predictive (all-pole) model they were able to classify 86% of the nasal murmur samples based on the frequency of the second pole.

Many studies in the 1970s and 80s refocused the parameters of inquiry to the region of the nasal-vowel junction. The focus on the boundary between the nasal murmur and the vowel was perhaps influenced by literature on adaptation, which suggested that auditory representations are enhanced over frequency regions where short-term amplitude changes take place, or by work in the perception of stop consonants suggesting time-varying properties for place of articulation. This region is marked by just such an increase in amplitude, from low-amplitude murmur into high-amplitude vocalism. Blumstein and Stevens (1979) investigated the shape of the spectrum at the region of the NV juncture (across various vowel environments), thus incorporating samples from the vowel onset transition. However, like the results of Fujimura (1962), Blumstein and Stevens failed to find invariant properties for [m] and [n] across all vowel contexts. Repp (1986) expanded the region of inquiry to include the portions of the murmur and the vocalic transition in his investigation of /NV/ acoustics across varying vocalic contexts. He concluded that there is no invariant spectral change property distinguishing [m] from [n] across different vowel contexts.

Kurowski and Blumstein (1987) extended the general method used by Repp (1986) of focusing on the dynamic aspect of the NV juncture. They showed that by transforming the units of analysis from Hertz to Bark (i.e., units based on the psychoacoustic characteristics of the human auditory system), general place of articulation characteristics for [m] and [n] in various vowel contexts emerge. There were (as expected) significant increases in spectral energy from the final few pulses of nasal

26

murmur into the initial pulses of the post-nasal vowel. The distribution of the increases in energy was place-specific, with an increase in the region of 5-7 Bark for [m] tokens and 11-14 Bark for [n]. These increases in energy at different Bark bands were found to hold across vowel contexts and also when the syllable-initial nasal was preceded by the fricative /s/.

The literature reviewed here has been concerned with locating those invariant aspects of the nasal acoustic signal that hold across varying environments. The present acoustic analysis of /m/, /n/ and /ŋ/ in syllable-onset position followed by the central vowel /a/ is focused on the region of the nasal-vowel juncture. The existing literature on the acoustics of nasal place investigates bilabial and alveolar places in syllable-onset position and largely neglects places posterior to the alveolar ridge. Given that the acoustic characteristics of nasal place at the NV juncture are similar to the acoustics of their homorganic oral counterparts (Cooper *et al.*, 1952), the acoustic features of the velar nasal at the NV juncture are predicted to be similar to those of /g/, with the post-nasal vowel onset exhibiting F2 values similar to those of alveolar tokens and F3 values lower than those of alveolars, revealing the so-called "velar pinch" (Stevens, 1998). Due to the lack of a front-cavity resonance for labials (Stevens, 1989), post-bilabial vowel onsets are predicted to have F2 values considerably lower in frequency than those of vowel onsets following either alveolar or velar closures.

2.2   Methods

2.2.1 Speech materials

Speech materials consisted of Filipino monosyllables read in isolation (i.e., not in a sentence frame). Randomized lists of three *target* syllables (that correspond to the stimuli used in subsequent perception studies), /ma/, /na/, and /ŋa/, were created with equal numbers of *distractor* syllables—/mi/, /ni/, /nu/, /ba/, /pa/, /ga/ to prevent repetition fatigue. Also included in the materials list were multiple repetitions of the nasal coda syllables /am/, /an/, and /aŋ/, along with the distractor syllables /im/, /in/, /un/, /ab/, /ap/, and /ag/. Each target syllable occurred 34 times. The entire list contained 400 monosyllables including nasal-onset distractors, and nasal-coda syllables. The syllables were written in Roman orthography using conventional Filipino spelling: "ma," "na," and "nga". The /a/ context was chosen for consistency with previous cross-language infant studies (Werker & Tees, 1984a; Best *et al.*, 1988). Additionally, /a/ is the most frequent post-nasal vocalic context in spontaneous Filipino speech (see Figure 1.2). The three target syllables are meaningful in Filipino and are given in Table 2.1 along with their English translations (Panganiban, 1972).

Table 2.1   Target Filipino speech materials and their English translations.

Token	Definition
ma	indicates the possessor of a quality
na	adv. now, already
nga	emphatic particle meaning "indeed"

2.2.2 Speakers

Three native speakers of standard Filipino were recorded reading the list of materials. All three speakers were female and affiliated with the University of British Columbia. Table 2.2 gives the background of the three speakers at the time of the recording.

Table 2.2. Background information for the three recorded speakers of Filipino.

Speaker	Age	Place of birth	Language(s) spoken at home	Time in Canada (years)
KL	20	Makati City	Filipino, Hokkien, English	3
DP	43	Manila	Filipino, Ilocano, English	9
FP	40	Manila	Filipino, Hokkien, English	10

Filipino was a first language for all three speakers. Two speakers also reported fluency in the Hokkien dialect of Cantonese[5], and one speaker grew up speaking Ilocano (Austronesian). None of the speakers reported any speech or hearing problems. Speakers were told that they were recording random Filipino monosyllables and none was aware of the precise nature of the study.

2.2.3 Setup and procedure

All three speakers were recorded individually in different recording sessions at the Interdisciplinary Speech Research Laboratory (ISRL) in the Department of Linguistics at the University of British Columbia. Recording sessions took place in a sound-proof booth (Industrial Acoustics Company, Inc.) at the ISRL. Speech materials were recorded through an external pre-amplifier directly onto the hard drive of an Apple

---

[5] There is a sizable ethnically-Chinese diaspora in the Philippines.

Macintosh G4 computer (housed in an adjacent room) using the SoundEdit audio program with an AKG condenser microphone approximately 8" from the speaker's lips. Materials were recorded in mono at a sampling rate of 44.1kHz and transferred to CD-R as *.wav files for analysis.

Speakers were asked to produce the monosyllables at a comfortable speaking rate. The experimenter sat in the recording booth with the speaker, prompting them, if necessary, to repeat tokens. Each recording session lasted approximately half an hour (with an additional half hour for explanation and completion of a language background questionnaire) for which the speakers were paid $20 CDN.

2.2.4 Analysis

Speech materials were analyzed using the Praat (v. 4.2.22 & 4.3.31) (Boersma & Weenink, 2005) software package for Microsoft Windows XP at the Infant Studies Centre, UBC. Both spectral and temporal characteristics of the target nasal-onset tokens were measured. The temporal measures were nasal murmur duration and vowel duration. The spectral measures were F1, F2, and F3 taken at the NV juncture. The higher-frequency spectral resonances and antiresonances of the nasal murmur, which undoubtedly contribute to the nasal place percept (Malécot, 1956; Recasens, 1983; Kurowski & Blumstein, 1993) but whose specific characteristics are highly variable in production (Fujimura, 1962), were not measured in the present analysis. Rather, the spectral measurements were focused on the nasal-vowel juncture, which were expected to show clear effects of place of articulation. Descriptions of these measurements are given below.

2.2.4.1 Temporal measurements

The duration of the entire /Na/ syllable was measured by placing cursors at the beginning and end of the token based on waveform and wide-band spectrographic displays. The onset of the syllable was taken as the first zero-crossing of the first periodic pulse of the nasal murmur. The offset of the nasal murmur and onset of the vowel was taken at the zero crossing of the first high-amplitude period of the post-nasal vowel. This location is defined as the "NV juncture" (Figure 2.1).
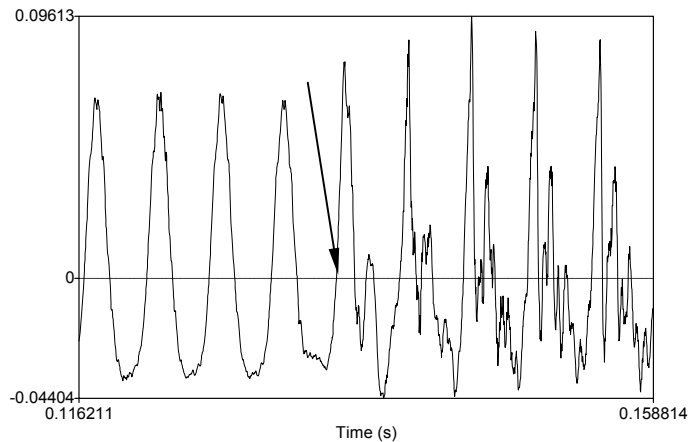
Figure 2.1 Four nasal murmur pulses and five vocalic pulses of sample [na] production. Arrow indicates the cursor location of the NV juncture.

The location of the NV juncture was confirmed by examining the spectrographic display which showed an abrupt increase in spectral energy at higher formants (F2 and F3) indicating the onset of vocalism. The nasal murmur is characterized by low-frequency, high-amplitude, pulsing in the waveform and high energy centered in the low frequency range and weak-energy higher resonances in the spectrographic display. The transition from nasal murmur to vocalic transition is marked (in most cases) by an abrupt increase in amplitude of the pulsing in the waveform. The placement of the nasal offset cursor in

the waveform display corresponded to the onset of high-energy formant patterning in the spectrographic display. An additional visual cue to the transition from nasal murmur to vocalism appeared in the waveform display, with the vocalic portion being marked by a more complex periodicity than the nasal murmur. In some cases, the increase in amplitude between the nasal murmur and vowel was gradual, thus precluding determination of the NV juncture based on the usual abrupt amplitude jump between the two regions. In those cases, the difference between the shape of the murmur and vowel waveforms (as seen in Figure 2.2) was the determining factor in cursor placement.
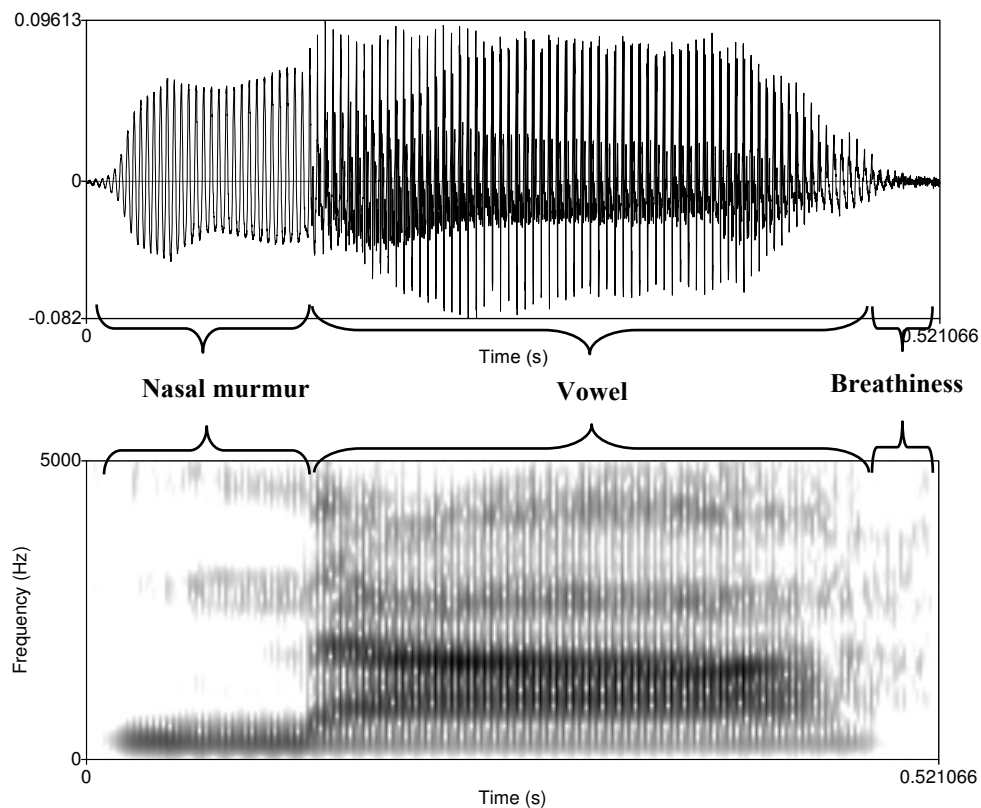
Figure 2.2. Waveform and wide-band spectrographic displays for a [na] token with vowel-offset breathiness.

As the offset of the vowel was often accompanied by high-frequency breathy noise, the offset cursor was placed at the offset of periodicity of the vowel based on both the wideband spectrographic display as well as the waveform display.

2.2.4.2 Spectral measurements

The formant measurements were based on FFT spectra calculated over a 10ms Hamming window centered at the NV juncture (as in Figure 2.1) and on a dB/Hz scale[6]. The 10ms window length was chosen to ensure that only frequency information local to the NV juncture (and not from nasal-murmur pulses or the adjacent vowel) was captured in the analysis. F1 was measured at the highest amplitude, non-fundamental low-frequency prominence in the display. In some instances, the higher formants (F2 and F3) did not correspond to a single prominent peak but rather to a group of highly prominent peaks. In these cases, the formant frequency was measured at the highest amplitude prominence.

---

[6] Recall that clear place characteristics emerged in Kurowski and Blumstein's (1987) analysis of nasal place using the Bark scale. Their analysis considered the change from nasal murmur into the post-nasal vowel. This change was characterized a an increase in energy in *ranges* along the Bark dimension. The current analysis differs from Kurowski and Blumstein's (1987) characterization by considering the place characteristics unique to the NV juncture, rather than a change-over-time measure.

FFT spectrum (10ms window), [ma]
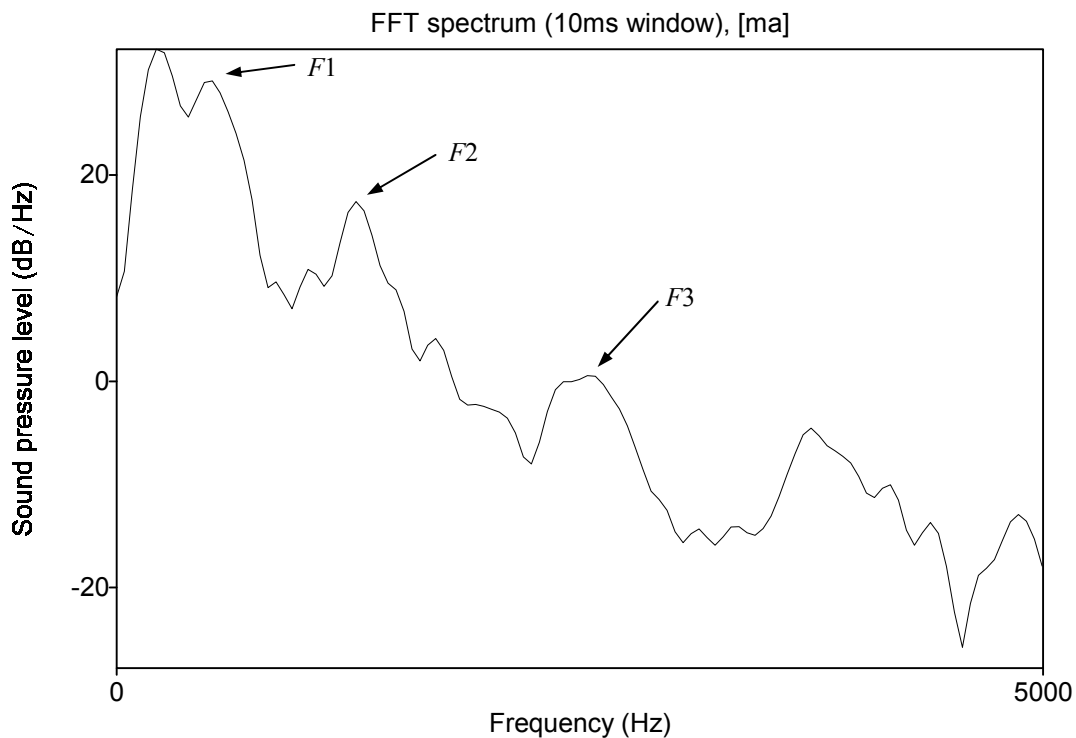
Figure 2.3. FFT (over a 10 ms window) spectrum of a [ma] token as measured at the NV juncture.

In ambiguous cases, the cursor placement in the FFT display was compared to the LPC

analysis (with a prediction order of 10) using the Burg algorithm with pre-emphasis

applied (as packaged in the Praat "Get Formant" command). The LPC comparison was

used in less than 5% of the token measurements. Formant frequencies (F1, F2, and F3)

were calculated at the NV juncture.

2.3   Results

A number of tokens were discarded due to mispronunciations, creakiness,

breathiness and other non-linguistic perturbations in the signal rendering them

unanalyzable. Approximately 25-30 tokens per nasal place per speaker were included in

the analysis. Large variances were found in the temporal measures. This wide spread in

34

murmur and vowel durations is attributed to across-speaker differences in speech rate.

As the frequencies of F2 and F3 (and their relationship) at the CV juncture are known to

cue place of articulation (Delattre *et al.*, 1955; Stevens, 1989; Stevens & Keyser, 1989),

these measurements were an *a priori* focus of interest in the analysis. Measurements

were submitted to a one-way ANOVA in SPSS (v.12.0 for Windows) with Place ([ma],

[na] or [ŋa]) as the fixed factor. The descriptive statistics of the measurements are given

in Table 2.3:

Table 2.3. **Descriptive statistics for measurement variables pooled across three speakers.**

Variable	Place	Mean	SD	SE	Min	Max
Murmur dur (ms)	[ma]	123.20	31.00	3.30	53.41	206.75
	[na]	120.63	32.93	3.47	46.37	213.33
	[ŋa]	134.93	44.19	4.71	38.17	238.84
Vowel dur (ms)	[ma]	305.08	91.86	9.79	130.28	447.95
	[na]	312.18	91.90	9.68	115.53	465.79
	[ŋa]	315.59	91.79	9.78	121.90	470.47
F1 (Hz)	[ma]	430.20	85.58	9.12	273.48	644.65
	[na]	365.83	93.03	9.80	196.58	730.11
	[ŋa]	354.81	94.95	10.12	200.45	686.64
F2 (Hz)	[ma]	1269.31	110.12	11.73	966.39	1527.59
	[na]	1800.48	147.86	15.58	1510.26	2284.48
	[ŋa]	1833.22	170.54	18.18	1517.15	2248.51

F3 (Hz)	[ma]	2778.43	232.62	24.79	2339.71	3261.25
	[na]	3013.45	211.38	22.28	2549.63	3456.81
	[ŋa]	2655.12	168.70	17.98	2291.08	3150.57

The analysis revealed significant effects of Token on murmur duration [$F_{(2,263)}=3.86$), $p<0.05$], F1 [$F_{(2,263)}=17.54$, $p<0.001$], F2 [$F_{(2,263)}=420.54$, $p<0.001$], F3 [$F_{(2,263)}=69.63$, $p<0.001$], and F3-F2 [$F_{(2,263)}=216.14$, $p<0.001$]. There was no effect of Token on vowel duration. These results were further explored in a series of multiple comparisons, the results of which are given in Table 2.4.

Table 2.4. Results of the multiple comparisons (Scheffe). Significant differences are given in bold.

Variable	Token	Comparison	Sig.
Murmur dur	[ma]	[na]	0.896
	[ma]	[ŋa]	0.105
	[na]	[ŋa]	**0.034**
F1	[ma]	[na]	**0.001**
	[ma]	[ŋa]	**0.001**
	[na]	[ŋa]	0.724
F2	[ma]	[na]	**0.001**
	[ma]	[ŋa]	**0.001**
	[na]	[ŋa]	0.323
F3	[ma]	[na]	**0.001**
	[ma]	[ŋa]	**0.001**
	[na]	[ŋa]	**0.001**

Interestingly, speakers produced longer [ŋ] murmurs than they did [m] or [n]. One possible explanation for the longer velar nasal is that articulatory implementation of the lowering of the posterior portion of the tongue body requires more time than does the

lowering of apex from the alveolar ridge involved in /n/ production (Hardcastle, 1973; Kuehn & Moll, 1976). Such an articulatory constraint may be operating in oral stops as well, where it is reported, that cross-linguistically, unaspirated velar stops show longer positive voice onset times than do other places of articulation (Lisker & Abramson, 1964).

Speakers' productions of [ma] had significantly higher F1 frequencies at the NV juncture than did their productions of [na] or [ŋa]. This result is consistent with productions of [b] in American English which have been reported to have a F1 onset frequency higher than that of either [d] or [g] productions. The higher F1 could be due to lower initial tongue body height during the production of the bilabial closure relative to its height during alveolar or velar nasal constrictions (Kewley-Port, 1982).

Individual formant values likely play less of a role in the perception of place than do combination metrics, such as the relation between F3 and F2 (e.g. Stevens, 1989). The scatterplot in Figure 2.4 captures this relation. These data are analyzed using a Discriminant Function Analysis (DFA) and a receiver operating curve (ROC).
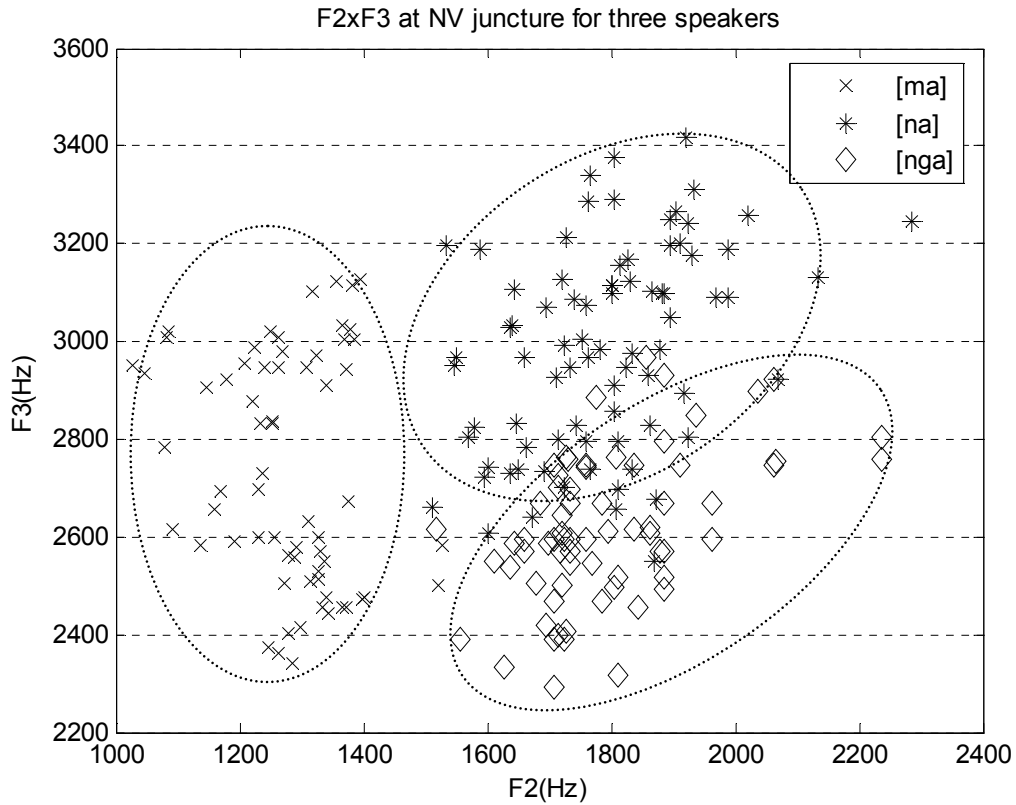
Figure 2.4. F2xF3 "clouds" (as measured at the NV juncture) representing three nasal places of articulation as spoken by three female speakers of Filipino.

A DFA[7] correctly classified 89.5% of the tokens. The analysis by token type is given in Table 2.5, which shows near-perfect classification of [ma] tokens. Most of the errors occurred in classifying [na] tokens as [ŋa]. Interestingly, the DFA does a better job predicting [ŋa] than [na], probably due the lower number of F3 [na] outliers.

---

[7] A discriminant function analysis essentially performs a regression for a categorial outcome. It builds a predictive model of group membership based on observed characteristics of each case. The analysis generates a set of discriminant functions based on linear combinations of the predictor variables that allow for the best separation between groups. The discriminant functions are determined from a sample of cases for which group membership is known. These functions are then applied to new cases with measurements for the predictor variables but unknown group membership (Based on the DFA description in SPSS v.12) .

Table 2.5. Confusion matrix produced by discriminant function analysis for classifying [ma], [na] and [ŋa] according to F2 and F3 at the NV juncture.

	Predicted group		
Token	**[ma]**	**[na]**	**[ŋa]**
[ma]	**97.7%**	0%	2.3%
	n=86	*n=0*	*n=2*
[na]	1.1%	**80.0%**	18.9%
	n=1	*n=72*	*n=17*
[ŋa]	1.1%	8.0%	**90.9%**
	n=1	*n=7*	*n=80*

A Receiver Operating Curve shows the trade-off between classifying a false positive and a false negative. A ROC analysis showed near-perfect separation between [ma] tokens and [na] and [ŋa] based F2 values. An F2 onset of 1525 Hz or lower correctly classifies 98.9% of the [ma] tokens as [ma]. There was more overlap in the ROC along the F3 dimension in classifying [ŋa]. For example, F3 values lower than 2895.73 Hz correctly classify 70% of the [ŋa], while incorrectly classifying 12% as [na] tokens.

2.4   Summary

The goals of this dissertation are to explore the role of acoustic and linguistic factors affecting the perception of nasal place of articulation in syllable-onset position. Crucial to this exploration is an understanding of the acoustic dimensions that differentiate the places of articulation in question. The present chapter examined three syllable-onset nasal places of articulation from Filipino along multiple temporal and

spectral dimensions. When spectral information is considered, information unique to the three places of articulation emerge. The [ma]-[na] distinction is differentiated by F1, F2, F3 and the F3-F2 difference, while the [na]-[ŋa] distinctions is characterized by F3, F3-F2 and a difference in the duration of the murmur.

To the extent that these place categories are not fully distinct, the combined F2 and F3 information as displayed in an F2xF3 space shows that [na] tokens overlap more with [ŋa] tokens than with [ma] tokens. Moreover, the discriminant function analysis of these data shows that the most misclassifications occur between [na] and [ŋa] tokens. These results are consistent with analyses of *oral* obstruent-vowel sequences from American English. Kewley-Port (1982) found a similar grouping of [da] and [ga] tokens in an F2xF3 space as measured at the onset of the vowel transition.

The acoustic "distance" between [ma], [na], and [ŋa] tokens is suggestive of a perceptual reflex. Given that the F2xF3 acoustic space is relevant for place perception (Stevens, 1989), we might predict that the perception of phonetic categories with more overlap in F2xF3 space would be measurably different from categories which show more separation in this acoustic space. This prediction is tested in Chapter 3, which presents the first in a series of experiments designed to assess the perception of /ma/, /na/ and /ŋa/ by adult listeners whose phonology reflects the three-way contrast (Filipino) and listeners for whom the /na/-/ŋa/ contrast is non-native (English).

CHAPTER 3: Adult cross-language perception of three nasal places of articulation

3.1   Introduction

The acoustics of [ma], [na] and [ŋa] presented in Chapter 2 revealed two basic points: acoustic measures differentiate the three places of articulation, but in F2xF3 space, known to be relevant for place perception (Liberman *et al.*, 1954; Delattre *et al.*, 1955; Larkey *et al.*, 1978; Stevens, 1989), [na] and [ŋa] tokens overlap resulting in imperfect separation of the two place categories in a discriminant analysis. This pattern, characterized by robust differentiation between [ma] and [na], and less robust differentiation between [na] and [ŋa], parallels the typological distribution described in Chapter 1, namely, when languages of the world contrast two nasal places of articulation in syllable-onset position, they most often do so at the bilabial and alveolar/dental places (Maddieson, 1984). Additionally, as noted in Chapter 1, the syllable-initial velar nasal occurred less frequently than either bilabial or alveolar nasals in the spontaneous speech of 60 Filipino speakers (section 1.5.2). The asymmetrical nature of the distribution of

syllable-initial nasals in Filipino, coupled with their asymmetric acoustic distribution in F2xF3 space, suggest that a parallel asymmetry will be evident in perception as well.

The present chapter asks first whether English-speaking adults, whose phonology does not contrast syllable-onset /n/-/ŋ/, show poor perception of the non-native contrast—a non-native effect that would be predicted by the reorganization hypothesis as discussed in section 1.1. Second, the research presented here asks whether Filipino speakers will exhibit a perceptual reflex of the acoustic nature of the contrasts described in Experiment 1 (Chapter 2), showing poorer perception of /n/-/ŋ/ than /m/-/n/ as predicted by the acoustic-perceptual salience hypothesis. Both of these questions are here assessed in a single cross-language perception experiment in which adult native speakers of either English or Filipino are presented with the contrasts [ma]-[na] and [na]-[ŋa] in a discrimination task.

As reviewed in Chapter 1, the general findings of cross-language speech perception (see Strange, 1995 for a full review) show that adult listeners' perception of non-native contrasts are influenced by native-language experience (Miyawaki *et al.*, 1975; Beddor & Strange, 1982; Werker & Tees, 1984a). For example, Werker and Tees (1984a) found that English-speaking adults were unable to discriminate reliably a dental-retroflex contrast from Hindi, while Hindi listeners did. This finding has been replicated with other non-native contrasts (Best *et al.*, 1988) as well as with synthetically-generated stimuli representing the dental-retroflex contrast (Werker & Lalonde, 1988). Such findings lead to the prediction, for the present experiment, that English listeners will perform according to English phonology, with accurate discrimination of [ma]-[na] and less accurate discrimination of [na]-[ŋa].

The predictions for the cross-language aspect of the present experiment are guided not only by the well-known findings of adult cross-language consonant perception, but also by two previous studies examining the perception of /m/, /n/ and /ŋ/ in syllable-onset position. Larkey, Wald, and Strange (1978) investigated the perception of synthetically generated onset and coda nasal place stimuli, [næ], [mæ], [ŋæ], [æn], [æm], [æŋ] along a continuum varying in the F2 and F3 frequencies of the pre- and post-nasal vowel. Using an AXB task, they found that English-speaking listeners' identification consistency and discrimination accuracy were inferior for the syllable-onset [næ]-[ŋæ] contrast compared to the coda [æn]-[æŋ] contrast. (In comparison, English listeners in that same study accurately identified and discriminated the [d]-[g] contrast in both syllable-onset and coda positions.) In a more recent cross-language AXB experiment using natural speech stimuli from Cantonese, Narayan (2004) found that English-speaking listeners were able to discriminate [ma]-[na] at rates above 90% correct, but were less than 70% accurate on the [na]-[ŋa] contrast. Cantonese-speaking listeners, for whom the velar nasal is phonemic in syllable-onset position, were able to successfully discriminate both the [ma]-[na] and [na]-[ŋa] contrasts[8]. In both the Larkey *et al.* (1977) and Narayan (2004) experiments, English-listeners' relatively poor performance in discriminating the non-native contrast was attributed to the lack of an onset /n/-/ŋ/ contrast in English. Similar results are expected in the present cross-language experiment.

---

[8] The syllable-onset /n/-/ŋ/ contrast is all but lost in the speech of Cantonese speakers from Hong Kong (Bauer & Benedict, 1997). The listeners in Narayan's (2004) study were from the Guangdong province of mainland China, where the contrast is retained.

The more crucial hypothesis advanced for the present experiment is that the acoustic similarity of [na] and [ŋa] may result in asymmetrical perception of the two contrasts by Filipino listeners. Filipino listeners are expected to discriminate both [ma]-[na] and [na]-[ŋa] at relatively high rates, but if the acoustic facts described in Chapter 2 affect listeners' perception regardless of language experience, Filipino listeners should perform less accurately on the [na]-[ŋa] contrast than on [ma]-[na] (as did the Cantonese speakers in Narayan, 2004). Such a result would suggest that the acoustics of syllable-initial /ŋ/—in comparison to /n/—can exert pressures on the accuracy of perception, which could potentially lead to asymmetries in its distribution (in the lexicon) and frequency (in spontaneous speech), as described in Chapter 1.

The purely acoustic-perceptual interpretation has to tempered, however, by the distributional facts noted in section 1.5.2. There is some evidence that phoneme frequency affects adult speech perception. Newman *et al.* (1999) show that lexical neighborhood effects on speech perception are subordinate to the effects of high frequency phonemes, while lower-frequency phonemes show effects of lexical neighborhoods. That is, listeners were biased towards identifying an ambiguous speech sound according to its phoneme frequency rather than its lexical neighborhood when the phoneme frequency was high. Results such as Newman *et al.*'s (1999) suggest that the effects of frequency cannot be separated from the effects of acoustic salience when assessing accuracy of speech perception.

Frequency effects aside, that acoustic similarity affects discriminability of speech contrasts has received attention in the vowel perception literature (Carlson *et al.*, 1979; Nord & Sventelius, 1979; Blandon & Lindblom, 1981), which has shown that acoustic

distances, calculated on the basis of various physical properties of the signal, are highly correlated to listeners' similarity judgments (with correlation coefficients generally between 0.80 and 0.90). Similarly, Krull's (1990) study of naturally produced Swedish voiced stops showed that size of acoustic differences corresponded to listener confusions in an identification task. For example, she found that [gɛ] and [ga] were close to [d] in F2xF3 space and this acoustic distance led to a corresponding perceptual similarity. More generally, Krull's multiple acoustic distance measures (including durations of stop bursts and formant frequencies of the following vowel) accurately predicted asymmetries in listener confusion. In continuing this line of reasoning, the present study hopes to show that the acoustic similarity of [na] and [ŋa] and the *dissimilarity* of [na] and [ma] have perceptual consequences for the Filipino listener.

In summary, the predictions for Experiment 2 are that 1) both English and Filipino listeners will perform well on the native [ma]-[na] contrast, 2) English listeners will show an effect of language experience with poor performance on the non-native [na]-[ŋa] contrast and 3) Filipino listeners will show a perceptual consequence of the acoustic distance of [na] and [ŋa] with somewhat depressed performance on the [na]-[ŋa] contrast relative to their performance on [ma]-[na].

3.2   Stimulus acoustics

Real-speech stimuli were selected from tokens recorded by native Filipino speaker KL, described in Chapter 2. Four tokens of each nasal place ([ma], [na] and [ŋa]) were selected based on their showing minimal within-category variation. The resulting 12 [Na] tokens had similar durations of nasal murmur and vowel as well as similar values for F1 at the onset of the post-nasal vowel transition. Variation in the vocalic portion (as

45

measured at the vowel midpoint) of the stimuli was avoided to the extent possible in order to preclude discrimination based on vowel quality. In addition, tokens were selected that differed minimally in the f0 of the nasal murmur and vowel. Table 3.1 shows the acoustic characteristics of the 12 stimuli. Formant measures were taken at vowel onset and midpoint as indicated. Fundamental frequency was averaged over a selected range, as described in Chapter 2. As the present, and all subsequent experiments are focussed on the place-distinguishing characteristics of the three nasals, statistical tests were conducted on the formant frequencies of the stimuli at the NV juncture.

Independent-samples t-tests revealed that the F2 onset frequencies were significantly higher than for [na] than for [ma] stimuli [$t(6)=-9.660$, $p<0.001$]. The difference in F2 onset between [na] and [ŋa] stimuli approached significance [$t(6)=2.401$, $p=0.053$], while F3 at vowel onset was significantly higher in [na] stimuli than in [ŋa] stimuli [$t(6)=3.971$, $p<0.007$]. No other acoustic measurements that were taken (i.e., temporal measures, or spectral measures taken at vowel midpoint) were significantly different between [na] and [ma] or [na] and [ŋa]. These results are very similar to the more general pattern of /Na/ acoustics of the productions of the three speakers of Filipino given in Chapter 2. The stimuli were then impressionistically assessed by three phonetically trained listeners to ensure that the tokens sounded similar in terms of overall pitch contour, duration and extra-linguistic factors such as breathiness or creakiness.

Table 3.1. **Onset-nasal stimuli acoustics.** Boxed averages represent critical spectral dimensions distinguishing nasal place of articulation.

	Token	Total dur	Nas dur	V dur	Total f0	V f0	F1 onset	F2 onset	F3 onset	F1 midpt	F2 midpt	F3 midpt
[ma]	1	498.21	123.35	374.86	213.73	208.97	420.43	1482.15	2543.87	1001.85	1604.48	2757.37
	2	512.70	138.18	374.52	217.72	213.44	419.87	1396.40	2451.06	1005.79	1472.00	2646.36
	3	479.61	131.53	348.08	216.73	215.04	424.66	1463.31	2654.70	974.53	1456.96	2685.25
	4	485.98	136.24	349.74	211.38	206.01	420.43	1431.59	2518.59	951.29	1419.35	2746.10
	Avg.	**494.13**	**132.33**	**361.80**	**214.89**	**210.87**	**421.35**	**1443.36**	**2542.06**	**983.37**	**1488.20**	**2708.77**
	SD	*14.59*	*6.60*	*14.90*	*2.89*	*4.13*	*2.22*	*37.62*	*84.70*	*25.50*	*80.62*	*52.29*
[na]	1	488.18	131.17	357.01	212.37	207.42	420.43	1760.22	2792.50	976.57	1507.43	2670.27
	2	511.17	127.62	383.55	210.47	207.67	346.78	1748.13	2624.33	915.57	1559.58	2793.01
	3	467.12	129.47	337.65	217.78	211.54	420.43	1734.94	2619.71	1001.85	1482.10	2653.07
	4	493.44	123.16	370.28	211.26	207.90	395.15	1861.34	2695.55	1001.85	1406.31	2619.71
	Avg.	**489.98**	**127.86**	**362.12**	**212.97**	**208.63**	**395.70**	**1776.16**	**2683.02**	**973.96**	**1488.86**	**2684.01**
	SD	*18.14*	*3.45*	*19.59*	*3.30*	*1.95*	*34.72*	*57.72*	*80.82*	*40.71*	*63.79*	*75.63*
[ŋa]	1	505.99	130.63	375.36	213.35	207.73	420.43	1684.38	2366.92	838.15	1429.43	2512.79
	2	497.98	123.50	374.67	211.92	209.13	395.15	1684.38	2543.87	976.57	1507.43	2594.43
	3	504.65	131.12	373.53	214.80	211.60	352.94	1739.03	2420.44	897.04	1470.94	2688.07
	4	461.16	124.98	336.18	NA	NA	398.27	1692.24	2475.16	1012.90	1595.19	2630.37
	Avg.	**492.45**	**127.56**	**364.94**	**213.36**	**209.49**	**391.70**	**1700.01**	**2451.59**	**931.17**	**1500.75**	**2606.42**
	SD	*21.15*	*3.88*	*19.18*	*1.44*	*1.96*	*28.18*	*26.28*	*75.74*	*78.65*	*70.57*	*73.38*

3.3   Methods

3.3.1   The AX discrimination task

The AX or 2-alternative forced-choice task is commonly used in cross-language speech

perception research (Beddor & Gottfried, 1995) and requires the participant to decide

whether two stimuli, which are presented successively, are the "same" or "different."

The choice of using an AX task rather than a task such as oddity, ABX, 4IAX, or AXB

was based on several factors. The most important factor had to do with interpretation of

the adult results from a developmental perspective. As the dissertation sets out to explore

the nature of nasal place perception across infancy and into adulthood, the adult

perception task needed to be functionally similar to the task used in the infant studies,

namely *habituation* (described in detail in Chapter 4), in order to compare results across

experiments. The habituation task essentially requires the infant to use the knowledge

acquired during the habituation phase in making a same-different judgment during the

test phase. The AX task was determined to be the most appropriate adult discrimination

task in both comparability with the infant task as well as efficiency in implementation

and interpretation of results.

3.3.2   Trials

AX (two-alternative forced choice) trials consisted of two types of *non-identity*

pairs, different pairs and within-category pairs, for both place-of-articulation contrasts

[ma]-[na] and [na]-[ŋa]. For the different (across-category AB) pairs, each of the four

tokens from each place category was paired with each of the four tokens from a

contrasting category (except [ma]-[ŋa] pairs were excluded). For example: $na_1ŋa_1$,

na$_2$ŋa$_1$, na$_3$ŋa$_1$, na$_4$ŋa$_1$; na$_1$ŋa$_2$, na$_2$ŋa$_2$…, and so on, for a total of 192 different pairs. Orders were counterbalanced in both different and within-category pairs. For within-category pairs, each token was paired with an acoustically different, but phonemically *same* variant (A$_x$A$_y$), for a total of 108 within-category pairs. Because within-category pairs did not consist of acoustically identical pairings, listeners would have to generalize across small acoustic differences in order to make categorical decisions.

The experiment was broken up into three blocks. Each block had 100 trials which consisted of one presentation of each possible pairing. The first eight trials of only the first block were "familiarization" trials meant to acquaint the listener to the sounds and the task. Responses to these trials were not analyzed. The inter-stimulus interval (ISI) was set at 1500ms as previous research has shown it to facilitate a "phonemic-level" of categorization of non-native phones (Pisoni, 1973; Werker & Tees, 1984b; Anderson *et al.*, 2003).

3.3.3   Participants

Fifteen native Canadian-English-speaking listeners and 13 native Filipino-speaking listeners were recruited from the Department of Psychology subject pool at the University of British Columbia. Listeners received either course credit or $10CDN for their participation.

Each participant was administered a language-background questionnaire before beginning the experiment. Canadian-English speakers were chosen to participate only if

they had not had any formal training in languages from East or Southeast Asia[9]. Similarly, only Filipino-speaking listeners who had spent their early years in the Philippines and continued to regularly speak Filipino with friends and family were chosen to participate. All the Filipino participants were immigrants to Canada and some reported fluency in other Austronesian languages of the Philippines (such as Ilocano, Bikolano, Cebuano, Kapampangan, etc.) as well as Spanish and French. The mean length of residence in Canada for the Filipino speakers was 4.92 (*SD*=3.42) years. None of the participants had received linguistics or phonetics training prior to the experiment.

3.3.4 Setup

The perception experiment was conducted using the DMDX (Forster & Forster, 2003) presentation software running on a Windows XP laptop computer in a quiet room at the Infant Studies Centre (UBC). Stimuli were presented through a Sony STR-DE197 amplifier over two free-field Bose (Model 101) speakers placed at 30° to the right and left of the centerline of the participant at 70±2 dB SPL[10]. Participants were tested individually and sat in front of the laptop computer which registered responses from clicks of right and left mouse buttons marked "S" (for same) and "D" (for different). Participants held the mouse on their lap with their right hand (the dominant hand of all the participants) and made responses with their index and middle fingers. Trials advanced automatically when the participant made a button press. There was a one-second delay between the offset of the button press and the onset of the next trial.

---

[9] Many South East Asian languages and some dialects of Cantonese exhibit the velar nasal in syllable-onset position. In order to avoid any listener bias, Canadian-English speakers who were familiar with other East and South East Asian languages were not chosen to participate.
[10] Stimuli were presented "free field," rather than the more common headphone presentation, for consistency with the anticipated infant experiments.

3.3.4.1 Orienting variables

Instructions for the English-speaking listeners were given on the computer screen in English. Participants were told that they would be hearing pairs of speech sounds and asked to press the "S" button if the two sounds were from the same category and "D" if they were from different categories. Participants were also asked to respond to the trials as quickly as possible. No other instructions were given to the listeners. The English instructions were translated into Filipino (by the native Filipino speaker who recorded the stimuli) and presented online to the Filipino-speaking participants. Although the Filipino speakers were necessarily bilingual with Canadian English, experimental instructions were presented in the listeners' native language in order to increase the likelihood that they performed the task according to L1 rather than L2 strategies.

Participants were told that the first 8 trials of the first block were familiarization trials meant to acquaint them with the speech sounds of the experiment and the response apparatus. Listeners did not receive any feedback during familiarization or the experimental phases. There was a short break after the familiarization in case the participant had any questions regarding the administration of the experiment. Listeners were given two short (2-5 minute) self-regulated breaks after the first and second blocks. The entire experiment lasted approximately 45 minutes.

3.4   Results

3.4.1   Proportion correct

Results from all 15 English listeners were included in the analysis. The results of one Filipino listener were discarded due to a misunderstanding of the instructions, leaving data from 12 Filipino speakers to be used in the analysis. Listeners' responses were converted to proportion correct scores for the [ma]-[na] and [na]-[ŋa] contrasts. These results show the predicted pattern of discrimination performance on both the native and non-native contrasts. English listeners accurately discriminated the native [ma]-[na] contrast at an average rate of 98.8% correct. Their performance on the non-native [na]-[ŋa] contrast, however, was slightly below chance at an average of 45.9% correct with considerably larger variances. The Filipino-speaking listeners on average showed the expected native level of performance, with an average of 98.8% correct on the [na]-[ma] contrast and 90.8% correct on [na]-[ŋa].

A 2x2x3 repeated-measures ANOVA was performed with the fixed factors Language (English and Filipino), Contrast ([na]-[ŋa] and [na]-[ma]) and Block (1, 2, and 3). There was a significant main effect of Language [$F(1,149)=65.57, p<0.001$], with Filipino listeners performing more accurately, overall, than English listeners. The main effect of Contrast was significant [$F(1,149)=122.35, p<0.001$] showing listeners performing better on the [ma]-[na] contrast than the [na]-[ŋa] contrast. The main effect of Block was not significant. The effects of Language and Contrast can only be interpreted in terms of their significant interaction [$F(1, 149)=66.18, p<0.001$].

Figure 3.1 shows the Language x Contrast interaction by plotting the average proportion correct on both contrasts for the two language groups. The interaction was

further probed with a series of one-way ANOVAs, which shows that Filipino listeners

were significantly more accurate than the English listeners on the [na]-[ŋa] contrast

[$F(1,78)=66.10$, $p<0.001$]. The language group difference for the [na]-[ma] contrast was

not significant. Taken together, the results show that the only difference between English

and Filipino listeners was in their discrimination of the [na]-[ŋa] contrast.

Proportion correct: English & Filipino listeners
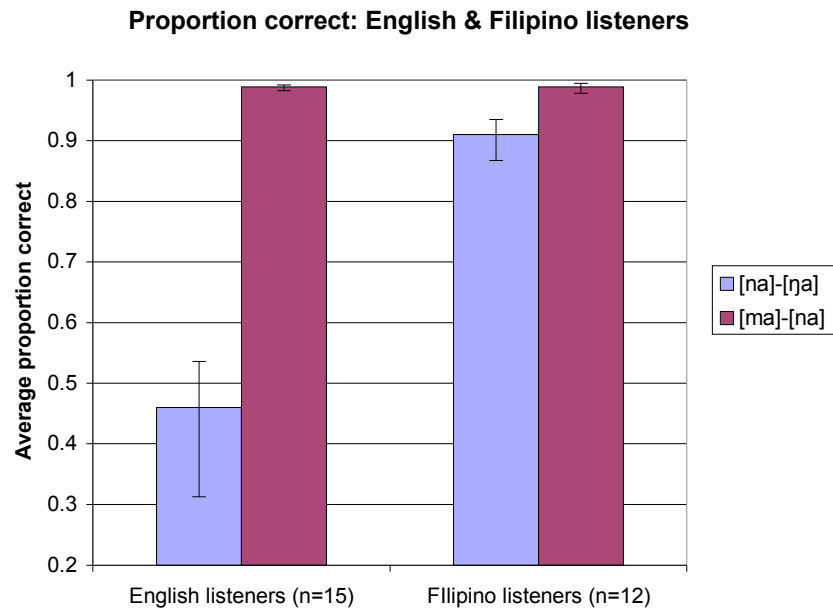
Figure 3.1 Average proportion correct on three contrasts according to language. Error bars represent standard error.

Within each language group, contrast differences were analyzed in a one-way

ANOVA. English listeners performed significantly better on the [ma]-[na] contrast than

on the [na]-[ŋa] contrast [$F(1,88)=126.12$, $p<0.001$]. Similarly, within the Filipino-

listener group performance was significantly better on [ma]-[na] than on [na]-[ŋa]

[$F(1,69)=24.62$, $p<0.001$].

53

3.4.2 Sensitivity (*d'*)

In addition to proportion correct scores, listeners' responses were converted to *d'*, a signal detection theoretic statistic that provides a bias-free measure of sensitivity to given contrasts (Swets, 1996; Macmillan & Creelman, 2005). In general, the higher the *d'* value the more sensitive the listener is to the contrast (i.e., the further apart in perceptual space are the two phonetic categories). The *d'* value is calculated from *hit* and *false-alarm* rates of the individual for a given contrast. The *d'* measure is useful in that it provides the researcher with a statistic based on listeners' perception of *both* same (within category) and different (between category) pairings. Table 3.2 shows the confusion matrix for any contrast with hit, false-alarm, correct-rejection, and miss rates defined according to given stimuli type and response.

Table 3.2 Confusion matrix used in calculating *d'* for each listener on every contrast.

Response→ Stimulus ↓	**Different**	**Same**
Different	Hits	Misses
Same	False-alarms	Correct rejections

Standard Signal Detection Theory assumes two types of AX discrimination designs, fixed and roving. In general, the roving design (or differencing model) is assumed when more than two stimulus classes are presented in the experiment. In the present task, listeners heard three classes, [ma], [na] and [ŋa], in any given block, therefore the roving design was used to calculate *d'* (see Macmillan & Creelman, 2005 for a full discussion of fixed versus roving models of discrimination). Macmillan & Creelman (2005) provide a *d'* table (based on false-alarm and hit rates) assuming the roving model. These *d'* values were found for each contrast in each block for every participant and used in the analysis.

As with the proportion correct results, listeners' *d'* scores were analyzed in a repeated-measures ANOVA with dependent variable *d'* and fixed factors of Language, Contrast and Block. There were significant main effects of Language [$F(1, 146)=51.89$, $p<0.001$] and Contrast [$F(1, 146)=134.53$, $p<0.001$] in the expected directions. There was a significant interaction between Language and Contrast [$F(1,146)=40.31$, $p<0.001$]. Interestingly the main effect of Block also reached significance [$F(2, 146)=4.11$, $p<0.05$] showing increased sensitivity to the contrasts as the experiment progressed. No interactions with Block reached significance.

Sensitivity : English & Filipino listeners
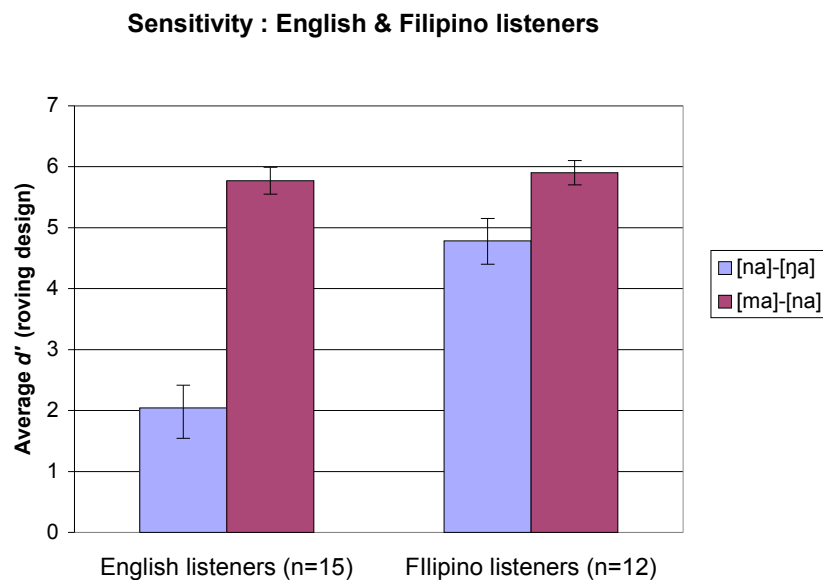
Figure 3.2: Average *d'* for three contrasts according to the native language of the listener. The sensitivity measure, *d'*, is calculated according to the roving design for the AX task (Macmillan and Creelman, 2005). Error bars represent standard error.

Figure 3.2 shows the interaction between Language and Contrast, which was further explored with one-way ANOVAs. Between language groups, Filipino listeners were more sensitive to the [na]-[ŋa] contrast than were English listeners [$F(1,78)=62.61$,

p<0.001]. Within the English group, listeners were more sensitive to the [ma]-[na] contrast than to the [na]-[ŋa] contrast [$F(1,88)=175.10$, *p*<0.001]. Within the Filipino group, listeners were similarly more sensitive to the [ma]-[na] contrast than to the [na]-[ŋa] contrast [$F(1,66)=12.60$, *p*=0.001].

The main effect of block on *d'* scores is plotted in Figure 3.3. The plot clearly shows listeners' increasing sensitivity (pooled across Contrast and Language background) as the experiment progressed.

Overall NV sensitivity
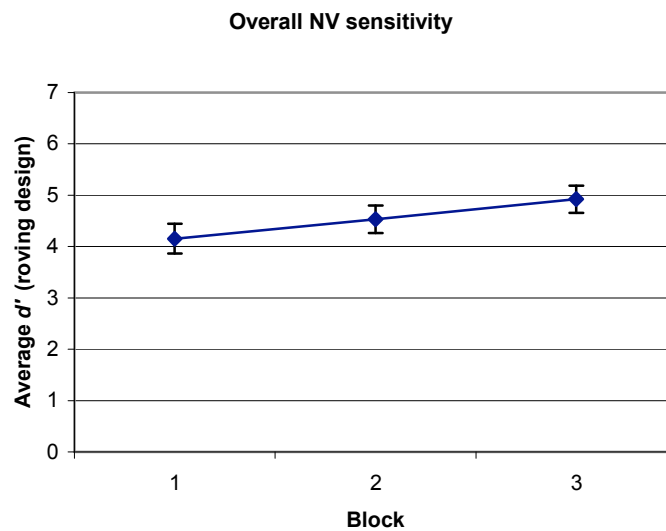
Figure 3.3. Listener sensitivity in each block pooling across contrasts and language groups. Error bars represent standard error.

3.5 Summary and discussion

Overall, the patterns of results from this real-speech cross-language AX experiment were consistent with results of previous investigation into English listeners' perception of synthetic nasal consonants (Larkey *et al.*, 1978) and English and Cantonese-listeners' perception of Cantonese nasals (Narayan, 2004). English-speaking listeners'

performance on the [na]-[ŋa] contrast was diminished relative to their performance on the native contrasts. Not surprisingly, the control group of Filipino-speaking listeners showed a perceptual pattern that reflected Filipino phonology, with highly accurate discrimination of both the [ma]-[na] and [na]-[ŋa] contrasts. These results are consistent with the prediction made in the introduction to the chapter, that listeners' perception of nasal place contrasts are influenced by their linguistic experience.

The more interesting result, however, was that both English and Filipino listeners were similar in their overall pattern of responses in that discrimination of the [na]-[ŋa] contrast was significantly poorer than that of the [ma]-[na] contrast. That Filipino listeners, for whom both contrasts are phonologically distinct, exhibited a pattern of relative discriminability parallel to that of the English listeners suggests that factors other than the nativeness of the contrast (in listener phonology) affects perception. A plausible explanation for this pattern of perception for Filipino listeners is that the [na]-[ŋa] contrast is perceptually less salient than the [ma]-[na] contrast. This reduced perceptual salience would most likely result from the different acoustic distances in F2xF3 space (which is known to play a significant role in consonant place perception) among the three contrasts shown in the results of Experiment 1 (Chapter 2): [na] tokens were further away from [ma] tokens than they were from [ŋa] tokens in F2xF3 space as measured at the NV juncture. The statistical confusion between [na] and [ŋa], as evidenced by the poorer separation of the two categories in a discriminant function analysis, is suggestive of the behavioral consequences in the present experiment. Here again, the interpretation of the reduced sensitivity to [na]-[ŋa] by Filipino listeners must also take into account the distributional facts of Filipino nasals described in section 1.5.2.

The sensitivity analysis revealed an interesting pattern not predicted in the introduction, namely that listeners' perception of both [ma]-[na] and [na]-[ŋa] improves with exposure. This learning, which was not evident in the proportion correct analysis, shows that the native and non-native listener can partially overcome perceptual biases with increased listening experience. Although the interaction between Contrast and Block did not reach significance in the sensitivity analysis, it is interesting that the Block effect seems to be driven by increasing sensitivity to the [na]-[ŋa] contrast.

Sensitivity across blocks: pooled results
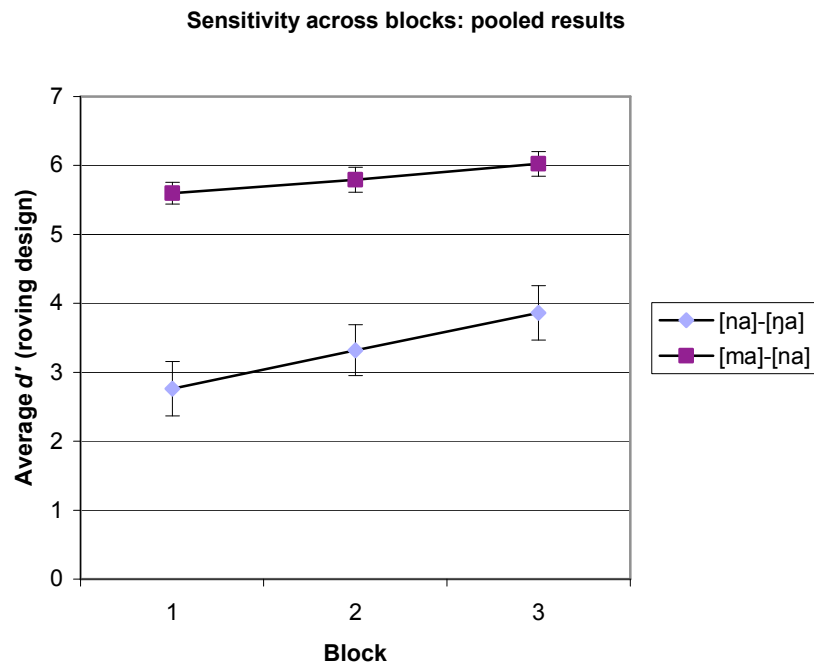
Figure 3.4 Listener sensitivity to the two nasal place contrasts, pooled across groups, according to block. Error bars represent standard error.

Figure 3.4 shows listener sensitivity results (pooled across language groups) to the two contrasts. Listeners increased their sensitivity to the [na]-[ŋa] contrast by roughly one *d'* unit. The corresponding increase in sensitivity to [ma]-[na] was not as dramatic even

58

though listeners could potentially further improve. Note that in the roving design for AX experiments in signal detection theory, perfect sensitivity is shown by a *d'* of 7.

As a group, English listeners performed at chance levels on the non-native [na]-[ŋa] contrast, yet the sensitivity of some individual listeners was similar to that of Filipino listeners. Figures 3.5 and 3.6 show the sensitivity measures for each listener in both the English and Filipino groups.

Individual listeners' sensitivity: English listeners
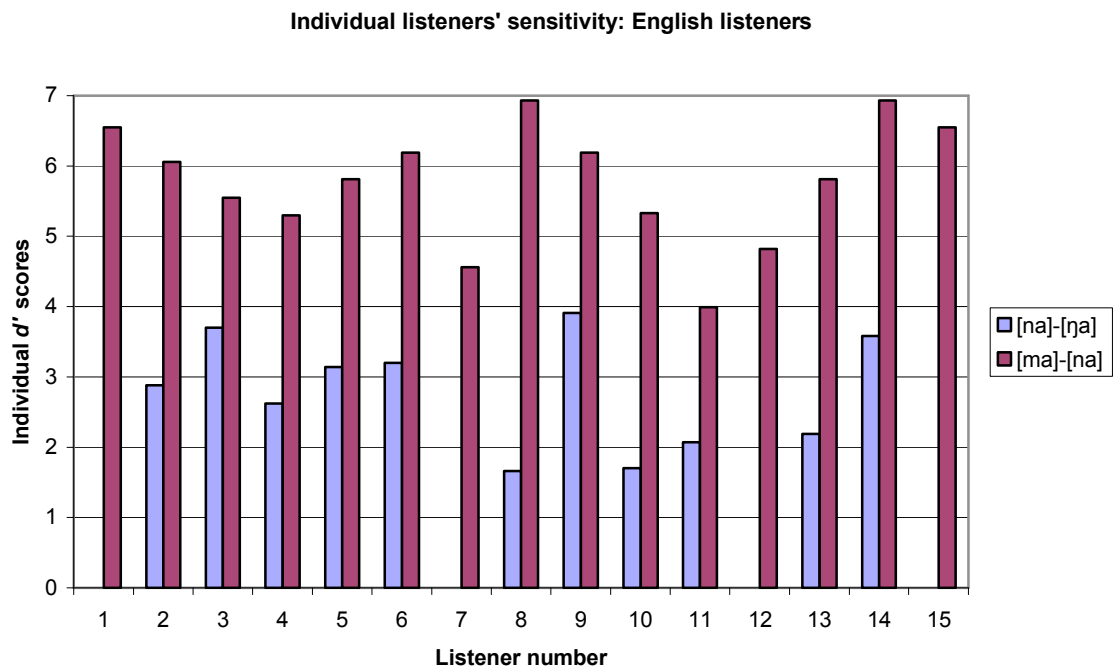
Figure 3.5. English listeners' individual *d'* scores. Listeners 1, 7, 12 and 15 had *d'* of 0 for the [na]-[ŋa] contrast.

Most English listeners were very sensitive to the [ma]-[na] contrast. The most variation was in the non-native contrast, where some listeners were completely insensitive (*d'*=0; listeners 1, 7, 12, and 15) while others were approaching a more native-like sensitivity

(listeners 3, 9 and 14). Filipino-speaking listeners exhibited individual patterns very similar to some members of the English group.

Individual listeners' sensitivity: Filipino listeners
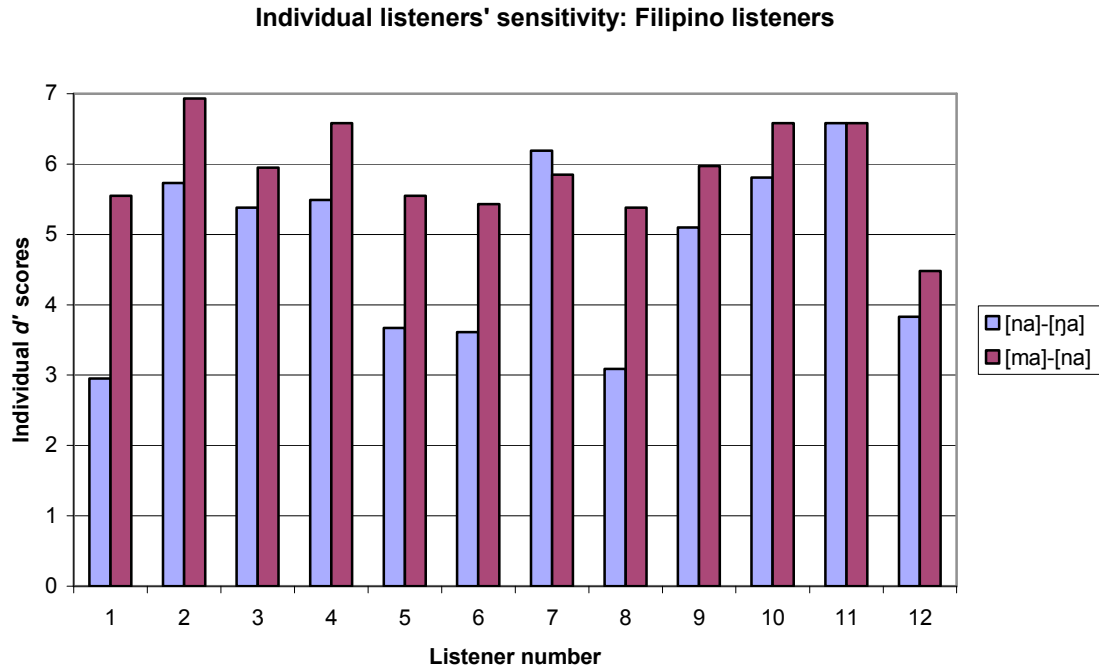
Figure 3.6. Filipino listeners' individual *d'* scores to the two nasal contrasts.

3.5.1 Conclusion

The present experiment establishes two findings: 1) that English listeners have difficulty discriminating the non-native nasal contrast [na]-[ŋa] (that is, the classic finding of non-native consonant perception) and 2) that Filipino listeners perform well on both nasal contrasts, but significantly better on [ma]-[na]. Together, these findings are interpreted as suggesting that the weak acoustic salience of the [na]-[ŋa] contrast results in correspondingly poorer discrimination accuracy relative to [ma]-[na]. The English listeners' poor performance on [na]-[ŋa]most likely results from *both* the influence of the non-native effect and the weak perceptual salience of the contrast. These results provide

60

motivation for an exploration into the development of these perceptual patterns given the well-known findings of non-native consonant perception in infancy discussed in Chapter 1. The experiments in the following chapters investigate whether the development of the perception of the nasal contrasts [ma]-[na] and [na]-[ŋa], which are perceived asymmetrically by adults, is affected by their relative acoustic distances and corresponding perceptual salience.

CHAPTER 4: English-hearing infants' perception of native and non-native nasal contrasts

4.1   Introduction

Experiment 2 (Chapter 3) revealed that, as expected on the basis of their phonological system, adult speakers of English have difficulty discriminating the non-native Filipino onset nasal contrast [na]-[ŋa]. These same listeners successfully discriminated the native contrast [ma]-[na]. The results from Experiment 2 support the now familiar pattern of mature listeners who are successful at discriminating acoustic differences that signal meaningful distinctions in their native phonology, and less successful at discriminating non-native contrasts (the non-native effect described in 1.2.1). Further, the results provide evidence that, apart from the non-native aspect of the contrast, the acoustic salience of the [na]-[ŋa] contrast affects its perception. Although the adult Filipino listeners in Experiment 2 accurately discriminated the [na]-[ŋa] contrast, their performance was significantly poorer than their performance on the [ma]-[na] contrast. Recall from the acoustic analysis in Chapter 2 that naturally produced [na] and [ŋa] tokens are closer together in F2xF3 space (measured at the NV juncture) than

are [ma] and [na]. I argue that this acoustic distance between [na] and [ŋa] and between [ma] and [na] may have contributed to the decreased level of discrimination of the [na]-[ŋa] contrast relative to the [ma]-[na] contrast by Filipino speakers (resulting in an *acoustic-perceptual salience* effect).

The goal of the present chapter is to chart the development of nasal perception in infancy with specific reference to the non-native effect observed in adult nasal consonant perception. The results of Experiment 2, specifically English-speaking adults' relatively poor discrimination of Filipino [na]-[ŋa] as well as Filipino adults' less accurate discrimination of native [na]-[ŋa] than [ma]-[na], provide crucial motivation for a cross-language investigation into infants' cross-language perception of the Filipino contrasts under the acoustic-perceptual salience and reorganization hypotheses discussed in sections 1.2 and 1.3. In every experimental report where adults have shown a non-native effect in speech perception, researchers have shown that infants' perception develops from reflecting language-general phonetic contrasts in early infancy, to language-specific contrasts towards the end of their first year (e.g., Werker & Tees, 1984a; Best *et al.*, 1988; Polka & Werker, 1994; Tsushima *et al.*, 1994, among others).   In this chapter, I conclude that perhaps reorganization, of the type evidenced by the results of Werker and Tees (1984a) does not fully account for the pattern of nasal discrimination observed in English-hearing infants between 4-12 months of age; rather, acoustic-perceptual salience, in combination with linguistic experience best accounts for infant performance on nasal contrasts.

I begin with the assumption that a reorganization effect will be observed in this set of experiments. At first blush, there is no obvious reason to expect nasal consonants,

as a class, to have a developmental-perception trajectory different from that of other consonants, for adults show non-native nasal perception patterns similar to those found with non-native oral contrasts. Therefore, one working hypothesis for the present set of experiments (Experiments 3 and 4) is that nasal consonant perception is not different from oral consonant perception, and the same developmental pattern will be observed. That is, English-hearing infants might show a non-native developmental effect by successfully discriminating the non-native [na]-[ŋa] contrast at 6-8 months, and failing to discriminate the contrast at 10-12 months. Infants at both ages groups are expected to discriminate the native [ma]-[na] contrast because it is phonological in English. However, a second hypothesis must also be considered given the results of the acoustic analysis (Experiment 1) and adult discrimination (Experiment 2): acoustic-perceptual salience may have an effect on infant performance, with the [ma]-[na] contrast being discriminated better than the [na]-[ŋa] contrast by young (as well as older English-hearing) infants.

Nasal consonants have not received the attention that oral stops (Werker *et al.*, 1981), fricatives (Nittrouer & Miller, 1997), velic and glottalic eggressives (Werker & Tees, 1984a; Best *et al.*, 1995), liquids (Tsushima *et al.*, 1994) and vowels (Polka & Werker, 1994) have received in the developmental speech perception literature. Hillenbrand's (1984) investigation of 5.5-6.5 month old infants' perception of bilabial and alveolar nasals across various vowels and speaker contexts was the first (and only) study to assess place within the nasal manner. In a conditioned head turn task, he found that infants, trained to discriminate the [ma]-[na] contrast as spoken by a male talker, were able to transfer their knowledge of nasal place to novel talker genders and vowel

contexts. Although Hillenbrand's study did not address the nature of *non-native* nasal consonant discrimination in infancy, his findings generally show that young English-hearing infants can discriminate the native [ma]-[na] contrast.

Experiment 3 is the first of four experiments I conducted to chart the development of infants' perception of nasal place contrasts. In this experiment, I tested English-hearing infants on the native [ma]-[na] contrast as spoken by a speaker of Filipino. Infants at two ages, 10-12 months and 6-8 months, are included in that experiment. As discussed in Chapter 1, these age ranges are informed by numerous studies in the developmental speech perception literature that suggest important changes occur in the infant's perceptual system in the second half of the first year of life. In this *control* experiment, infants are expected to discriminate the contrast; failure to discriminate would be surprising and would likely point towards a general difficulty for infants to discriminate place in nasals (or to inappropriateness of the task; see below). This prediction is not only guided by Hillenbrand's (1984) results, but also by the more general finding for consonant discrimination in infancy (Werker & Tees, 1984a; Tsushima *et al.*, 1994; Polka *et al.*, 2001; Anderson *et al.*, 2003).

In Experiment 4, English-hearing infants at 10-12 and 6-8 months of age are presented with the non-native [na]-[ŋa] contrast. As noted above, the two hypotheses described in Chapter 1 make differing predictions for perception of the non-native contrast. Under the reorganization hypothesis, which follows from the well-known findings in cross-language infant speech perception, the younger age group is expected to discriminate the non-native contrast, while the older group is predicted to fail. The acoustic-perceptual salience hypothesis predicts that the [na]-[ŋa] contrasts will not be

reliably discriminated by infants because of it being less salient than the [ma]-[na] contrast. The results of Experiment 4 suggest that both groups of infants (at 10-12 and 6-8 months) fail to discriminate the non-native contrast, consistent with an acoustic-perceptual salience interpretation. Before abandoning the reorganization hypothesis, however, even younger infants are tested on the native and non-native contrasts. In Experiment 5, I present 4-5 month old English-hearing infants with the native and non-native nasal contrasts.

All of the infant studies presented in this dissertation were conducted using the visual habituation paradigm. Researchers studying infant speech discrimination have successfully implemented this technique for assessing phonetic discrimination in infancy in numerous previous studies. The technique is discussed in detail below.

4.2   Experiment 3: Native nasal-contrast perception at 10-12 and 6-8 months of age

4.2.1   Methods

4.2.1.1 Assessing infants' discrimination of speech sounds: Visual habituation

I measured infants' discrimination of the native [ma]-[na] and non-native [na]-[ŋa] contrasts using the visual habituation (VH) paradigm. I used this procedure because of its applicability to a wide range of ages of infants as well as the ease and speed of its execution (Werker, Shi, Desjardins, Pegg, Polka, and Patterson, 1998). Habituation techniques have been used with infants as young as 7 weeks (Pegg *et al.*, 1992) and

upwards of 14 months (Fennell & Werker, 2003). Additionally, VH can be conducted

with only one experimenter and takes approximately 8 minutes to run.[11]

The general phenomenon of habituation is the progressive diminution of a

behavioral response with the repetition of a stimulus. Upon habituation, the presentation

of a new stimulus has the potential to evoke a recovery in behavioral response or

"dishabituation." Its application to assessing perception in infancy was first explored by

Horowitz (1975), who found that infants show a reliable increase in visual fixation times

when a background auditory stimulus is changed.

The general procedure of the visual habituation paradigm in speech perception

begins with the presentation of auditory stimuli, which is contingent upon the infant's

visual fixations on a visual pattern. Habituation to the auditory "background" stimuli is

indexed by a decrement in visual fixation to some pre-determined criterion, at which

point a *shift* to a new set of auditory stimuli occurs. Discrimination between the

background and shift stimuli is determined by a significant postshift recovery in visual

fixation time (to the same visual target). Best *et al.* (1988) were the first to use this basic

procedure of visual habituation (VH) in their assessment of the development of non-

native perception of Zulu clicks by English-hearing infants. Subsequently, Pegg *et al.*

(1992) and Polka & Werker (1994) have adapted the procedure in their studies of infant-

versus adult-directed speech and English-hearing infants' perception of a German vowel

contrast, respectively. In Best's *et al.* study, infants were habituated to background

stimuli and their looking time (LT) to a checkerboard image was measured on-line by an

observer. When the subject's LT decreased by 50% or more on two consecutive

---

[11] The conditioned head turn task, on the other hand, requires at least two experimenters to implement and takes at least 20 minutes to run.

habituation trials (habituation), relative to the LT for the longest two habituation trials, the *test* (or shift) stimuli were played. These stimuli were played, and their LT recorded, until the infants habituated to them as well, at which point the testing session concluded.

The procedure was implemented in a slightly different fashion at Dr. Janet Werker's Infant Studies Centre (ISC) in the 1990s (Pegg *et al.*, 1992; Polka & Werker, 1994). In both the Pegg *et al.* (1992) and Polka & Werker (1994) studies, infants were presented with *two* test trials after habituation. After the presentation of two test trials an additional two habituation trials were presented. These *re-habituation* trials served as a within-subjects control for random increases in LT not directly resulting from the stimulus change. If the LT to the re-habituation trials were similar to the pre-shift LTs, then the increased LT observed during the test trials could be attributed to the change in stimulus rather than random variation in LT.

It is important here to note the difference between VH and the other commonly used method for assessing discrimination, namely the Conditioned Head Turn (CHT). CHT was used in many cross-language infant speech perception studies such as those conducted by Werker and Tees (1984a), Kuhl *et al.* (1992), and Polka and Werker (1994) among others. The method has been described in detail by Werker, Shi, Desjardins, Pegg, Polka, and Patterson, (1998) and involves conditioning the infant to turn her head towards an animated toy whenever she detects a change in the background stimuli. One advantage of the CHT over VH is that individual infant performance can be assessed and sensitivity (in terms of a signal theoretic statistic) measured. This method is falling out of favor in many infant speech perception labs due to the time involved in running individual infants (over 30 minutes) as well as the number of researchers required

(typically 2). The theoretical implications of employing VH rather than CHT are discussed in the concluding chapter.

4.2.1.2 VH in the present study

The VH paradigm as implemented in the present study differs from previous VH studies in various ways. Looking time was calculated over groups of three (rather than two) trials and the criterion for habituation was set at a decrement in LT of 40% (rather than 50%) relative to the LT for the longest three consecutive trials. The 40% criterion was chosen in consultation with Dr. Werker and colleagues at the ISC who also study discrimination using VH. The criterion was an educated compromise to the 50% standard because infants from a wide age range would be tested (4-12 months). As younger infants need more time to habituate and older infants less time (Hunter & Ames, 1988), 40% seemed appropriate for reducing attrition while maintaining a strict criterion for habituation.

Another way VH differed in the present study from other VH studies was in the presentation of test trials. In the present study the infant received two post-habituation test trials, one "same" trial in which the individual tokens were the same as the tokens in the habituation trial but in a different order, and one "change" trial in which tokens that were phonetically different from the habituation tokens were presented. The "same" and "change" trials were counterbalanced in presentation. This design is most similar to the implementation of the *switch* procedure used by Werker and colleagues to assess word-object associations (Werker, Cohen *et al.*, 1998), and allows the researcher to know if the

infant shows a novelty or familiarity preference after habituation[12]. This method has

recently received theoretical support in the infant development literature (Cohen, 2004).

4.2.2 Participants

Infants were recruited from a database at the ISC. The database contained names

and phone numbers of parents who had earlier given consent for solicitation to participate

in infant studies[13]. Data from 32 infants[14] were analyzed in the study, 16 infants aged 10-

12 months (mean=303 days, range=291 to 332 days) and 16 aged 6-8 months (mean=211

days, range=177 to 241 days). An additional 10 babies participated in the study, but had

to be excluded from analysis due to crying or fussiness (*n=5*) or parental interference

(*n=2*). Three babies were excluded due to the strict habituation criterion described

below, whereby infants were included in the analysis *only* if they habituated in more than

6 trials. All the infants who participated were being exposed to at least 90% Canadian

English according to parental reports.

Infant visits generally lasted half an hour, which included time for greeting and

explanation of the procedure to the parent or caregiver, the experiment itself, and

debriefing. Infants received an "Infant Scientist" t-shirt and a faux "Infant Scientist"

diploma certificate for their participation in the study.

---

[12] There has been some debate in the infant behavior circles regarding the misuses of the habituation paradigm. Specifically, many researchers compare LT of test trials to LTs for habituation trials. Looking times to criterion trials are artificially low and may skew the interpretation. Cohen (2004) recommends that researchers include *both* a familiar trial and a test trial during the test phase. This is precisely what is done in the present set of infant studies.
[13] ISC staff would regularly visit new mothers at the BC Women and Children's hospital and describe the activities of the center. Additionally many parents would voluntarily respond to public service announcements placed by the center at events like "Movies for Mommies," or advertisements in local parenting newspapers.
[14] Infant gender was not counterbalanced in either age group.

4.2.3 Auditory stimuli

The stimuli for the infant study were identical to those used in the adult study (Experiment 2). Only alveolar ([na]) and bilabial ([ma]) tokens were presented to the infants in Experiment 3. Trials (either habituation or test) were approximately 14 seconds long. Two trial sets were assembled for both [na] and [ma] stimuli. Each trial consisted of 3 acoustically different tokens of the stimulus repeated 3 times, for a total of 9 tokens. The order of the tokens within each trial was randomized. The order of the tokens within the second trial (for the same phonetic stimulus) was randomized but consisted of the *same* tokens as in the other trial. The organization of each trial type is laid out in Table 4.1.

Table 4.1. Assembly of each 14 sec trial used in the native-nasal contrast study.

[na]		[ma]	
Trial type 1	Trial type 2	Trial type 1	Trial type 2
na_1	na_2	ma_2	ma_3
na_3	na_3	ma_3	ma_2
na_1	na_1	ma_2	ma_3
na_2	na_1	ma_3	ma_1
na_3	na_3	ma_1	ma_2
na_1	na_2	ma_3	ma_1
na_2	na_3	ma_2	ma_2
na_2	na_1	ma_1	ma_3
na_3	na_2	ma_1	ma_1

The inter-stimulus interval (ISI) was approximately 1000 ms.

4.2.4 Setup and apparatus

All infant studies were conducted in a quiet room at the Infant Studies Centre in the Department of Psychology at the University of British Columbia. The testing room measured approximately 9'x 8'. Infants sat on their caregiver's lap, who was seated in a chair approximately 4' from a 27" color TV monitor. The monitor was approximately 3.5' off the ground. Below the monitor were an amplifier and two Bose 101 speakers which were approximately 30° to the left and right of centerline of the caregiver's chair. A small JVC video camera (GX-N7UT) stood on a tripod directly below (approximately 1') the TV monitor.

The amplifier was connected to an Apple Macintosh G4 (OSX) computer in the adjacent control room. The *Habit 2002* software program (Cohen *et al.*, 2004) was used to present both visual and auditory stimuli, as well as record the infant's on-line LT. The video camera provided a closed-circuit broadcast of the infant's face to another monitor in the control room. The observer would watch the infant on the control room monitor and record "looks" to the TV monitor in the testing room on the Macintosh computer.

The entire testing room was draped in black cloth, including the periphery of the TV monitor. This draping occluded the camera, amplifier and wires or other ephemera which could potentially distract the infant from looking at the TV monitor. The lens of the digital video camera peeked out of a 2.5" hole in the black curtain. The caregiver wore noise-canceling headphones and listened to folk music during the entirety of the experiment. A mirror was strategically placed behind the caregiver, to allow the observer in the control room to monitor the visual stimulus. This reflected image was captured on tape by the video recorder and used by the observer in calculating *off*-line LTs.

4.2.5 Procedure

Every infant was given a unique code that was used for identification in post-study analysis. This code was registered onto the videotape used for recording the infant's behavior during the task and also noted in the *Habit* program. A dry run of the experiment was usually conducted in order to ensure that the auditory stimuli were presented at the proper volume. A Radio Shack decibel meter was used to adjust the free-field audio presentation to 70 ±2 dB SPL.

After the infant and caregiver were taken into the testing room the experimenter returned to the control room to begin administering the audio and video stimuli. On-line LT was recorded by depressing a designated key on the computer keyboard whenever the experimenter determined that the infant was looking at the TV monitor image.[15] These looks were recorded by the *Habit* program and used to calculate habituation. Habituation was determined by calculating the infant's LT to the unbounded image of a red and black checkerboard (Figure 4.1b) over a window of three trials. The program calculated the LT for successive three-trial windows and once a window fell to 60% of the looking time (a decrement of 40%) of the longest three-trial window, habituation was achieved. With this non-overlapping window, habituation occurred at 6, 9, 12, 15, 18, 21, or 24 trials.

Given the setup of the *Habit* program, a random depression in looking time during a given trial (due to dropping a toy or sneezing, for example) in the habituation period could potentially trigger the program to start the *test* trials. For example, if the infant sneezed during a habituation trial and as such closed her eyes for a long period prior to the sneeze, the LT could deceptively fall to 60% of the looking time of the longest

---

[15] The experimenter practiced on-line coding, under the supervision of a trained expert at the ISC, numerous times before running any babies through this study.

window of three trials, signaling habituation. For this reason, an *a priori* decision was made to exclude from the analysis any infants who habituated to the background stimulus in less than 9 trials (that is, infants who habituated in 6 trials). This ensured that infants minimally heard the background stimulus for approximately 2 minutes. If an infant's LT to any given trial (habituation or test) was less than 1 second, the trial was repeated. Infants were excluded from analysis if their LT to any test trial was under 1 second. Additionally, some infants were excluded from analysis due to their failure to habituate in the maximum number of habituation trials (24).

Trials were separated with a visual "attention getter" consisting of an animated circular sunburst against contrasting backgrounds. The attention getter was not accompanied by an auditory stimulus. When the infant re-fixated on the attention getter the experimenter would begin the next trial. Infants were either habituated to [ma] and tested on [na] or vice versa.
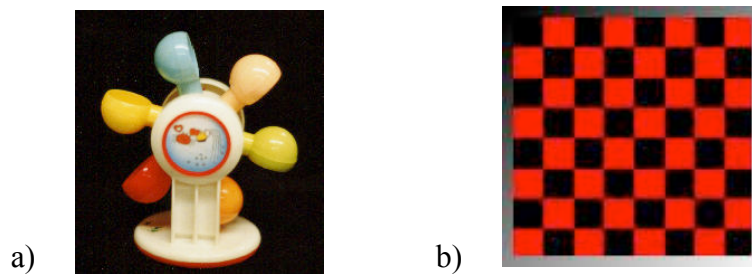
Figure 4.1. Visual stimulus for a) pretest and posttest and b) habituation and test trials.

The experiment began and ended with an assessment of the infant's interest in the task in general, a so-called *pretest* and *posttest.* The visual stimulus for these tests was an animated spinning waterwheel (Figure 4.1a). This image was accompanied by short tokens of a randomly rising and falling sine wave tone. Looking time to the pre- and

posttest was used to gauge the infants overall participation in the task. If infants' LTs to the posttest was significantly shorter than their LT to the pretest, then the reliability of the habituation would be in question for it indicates a lack of interest or participation in the task in general.

4.2.5.1 Off-line coding

For each infant who completed the study (i.e., completed the posttest without crying, habituated in greater than 6 and less than 24 trials, and did not repeat test trials) their accompanying video tape was digitized to QuickTime (Apple Inc.) format using the video editing software FinalCut Pro for the Macintosh computer. These digitized files were then coded *off*-line by the experimenter who was blind to which test trial was a *change* or *same* trial. Using a specially written script for the Excel (MS Office) spreadsheet program, the pretest, the last three habituation trials, the two test trials and the posttest were coded by the experimenter. The digitization process rendered the video file at a resolution of 30 frames per second. The experimenter proceeded through the video file, frame-by-frame, logging infant looks and non-looks. This off-line coding was thought to be more reliable than the on-line coding conducted during the experiment, whose only use was to determine habituation. Only data from the off-line coding were used in the analyses.

4.2.6 Results

Infants' average LT to same and change trials according to age group are plotted in Figure 4.2. The plot shows that both groups of infants looked longer to change trials than same trials.

English-hearing infants at 10-12 and 6-8 mos: [ma]-[na]
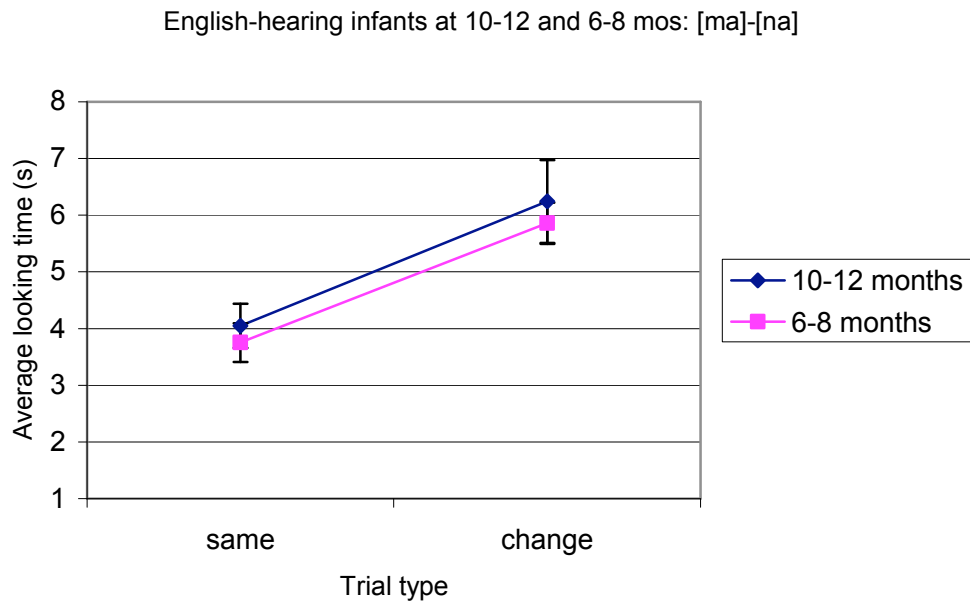
Figure 4.2. English-hearing infants' (at 10-12 and 6-8 months) looking time to same vs. change trials for the native [ma]-[na] contrast.

The LTs for each infant were computed for 1) the pre- and posttests and 2) the *same* and *change* trials. A 2 (Age group: 10-12 months vs. 6-8 months) x 2 (Test: pretest vs. posttest) repeated-measures ANOVA was conducted to determine if there were significant age-related differences in the infants' interest in the task. There was no significant effect of Age group or Test [$Mean_{PRETEST}$ =12.77s, $Mean_{POSTTEST}$=12.59s], suggesting that infants maintained interest throughout the entirety of the experiment.

Discrimination of the [ma]-[na] contrast was analyzed in a 2 (Age group: 10-12 months vs. 6-8 months) x 2 (Habituating stimuli: [ma] vs. [na]) x 2 (Test trials: same vs.

change) repeated-measures ANOVA. There was a significant main effect of Test trials showing that infants looked significantly longer to *change* trials than they did to *same* trials [$F(1, 67)=19.31$, $p<0.001$; Mean$_{SAME}$=3.93s, Mean$_{CHANGE}$=6.02s]. Across both age groups 27 of the 32 infants tested looked longer to change trials than same trials. There were no significant effects of either age group or habituating stimuli. No interaction was significant.

4.2.7   Discussion

That infants looked significantly longer to *change* trials than to *same* trials indicates that English-hearing infants in the second half of their first year can discriminate the native nasal-stop contrast [ma]-[na]. These results are consistent with the one other investigation into English-hearing infants' perception of this contrast (Hillenbrand, 1984) but is also consistent with both the reorganization and acoustic-perceptual salience hypotheses. Infants' success in discriminating the native [ma]-[na] contrast is consistent with the general finding for the development of consonant perception, whereby infants do not show a developmental progression in the discrimination of native contrasts. The [ma]-[na] contrast is relatively acoustically salient, thus allowing for accurate discrimination in infancy under the acoustic-perceptual salience hypothesis. Importantly, the present findings show that the VH procedure is sensitive enough to reveal discrimination of place of articulation within the nasal manner. As Hillenbrand's study employed the conditioned head turn paradigm, a finding inconsistent with his could have potentially been due to methodological factors.

As discussed in Chapter 1, based on the well-known pattern of the development of speech perception, which proceeds from a language-general to language-specific

reorganization of speech perception at 10-12 months, we might predict that older

English-hearing infants (10-12 months) will have difficulty discriminating a non-native

nasal place contrast. If they are unable to discriminate the Filipino [na]-[ŋa] contrast,

then we can attribute the failure to phonological factors rather than a general

imperceptibility of nasal place in infancy, for the results of Experiment 3 would suggest

otherwise.

Experiment 4 explores English-hearing infants' perception of the non-native nasal

contrast. As in Experiment 3, the performance of two age groups of infants is tested in

order to capture the non-native effect in the development of nasal consonant perception.

4.3   Experiment 4: Non-native nasal contrast perception at 10-12 and 6-8 months of age

4.3.1   Introduction

The results of Experiment 3 show that English-hearing infants at both 6-8 and 10-12

months of age successfully discriminate the native and acoustically salient [ma]-[na]

contrast. In Experiment 4, I present English-hearing infants with the non-native [na]-[ŋa]

contrast from Filipino. Under the reorganization hypothesis older infants are predicted to

be poor at discriminating the non-native contrast, while younger infants are predicted to

successfully make the discrimination. The acoustic-perceptual salience hypothesis

predicts that the weak acoustic salience of the [na]-[ŋa] contrast will affect its

discrimination in infancy.

4.3.2   Methods

The procedure and setup for the present experiment were identical to those in Experiments 2 and 3. The auditory stimuli, which were the same as those used in Experiment 1, were assembled in a fashion similar to that described in 4.2.3 above. The trials consisted of [na] and [ŋa] tokens.

4.3.3   Participants

As in Experiment 3, infants were recruited from the database at the ISC. Data from 32 infants (16 10-12 month olds and 16 6-8 month olds) were analyzed. In order to achieve this set of 32, 51 infants were tested. Among the 10-12 month old infants, 14 babies were not included in the analysis for the following reasons: crying or fussiness ($n=5$), parental or care-giver interference ($n=1$), experimental error ($n=1$), habituating in 6 trials ($n=6$), and failure to habituate ($n=1$). Among the 6-8 month old infants, 4 babies were not included due to parental or caregiver interference ($n=1$), repeated test trials ($n=1$), habituating in 6 trials ($n=1$), and failure to habituate ($n=1$). As in Experiment 3, all infants who participated were being exposed to at least 90% Canadian English according to parental reports.

4.3.4   Results

Infants' attention to the task was assessed in a 2 (Age group: 10-12 months vs. 6-8 months) x 2 (Test: Pretest vs. Posttest) repeated-measures ANOVA. There were no significant effects of Age group , Test, or their interaction, confirming that infants maintained interest in the task over the course of the experiment [$Mean_{PRETEST}=12.74s$,

Mean$_{POSTTEST}$=12.49s]. Infants' average LTs are plotted in Figure 4.3 according to age group.

English-hearing infants, 10-12 and 6-8 months: [na]-[ŋa]
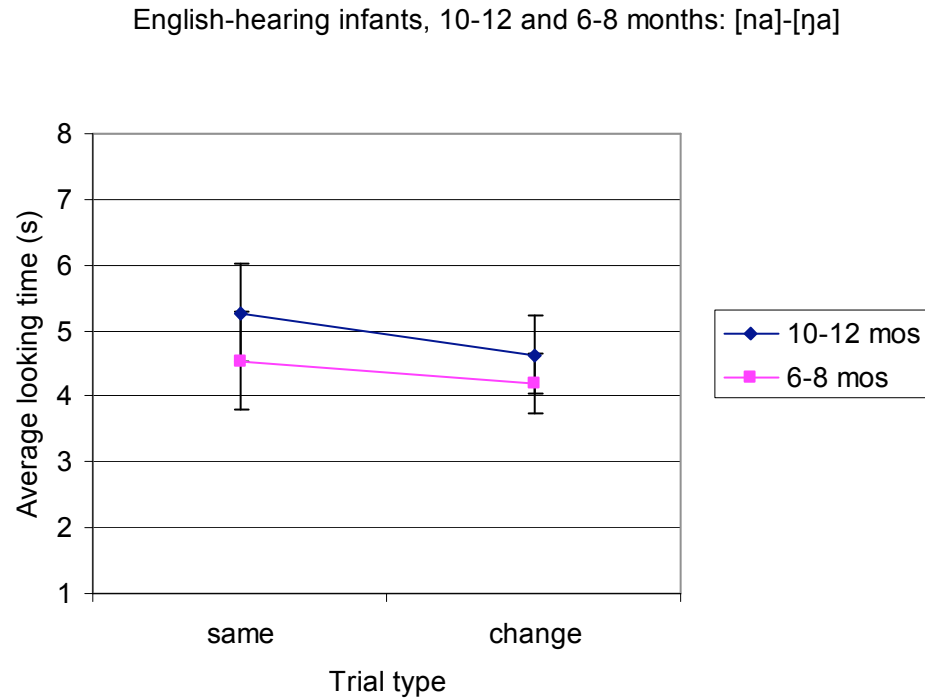
Figure 4.3. English-hearing infants' (at 10-12 and 6-8 months) looking time to same vs. change trials for the non-native [ma]-[ŋa] contrast.

In order to assess infants' discrimination of the non-native nasal contrast, infants' LTs to same and change trials were analyzed in a 2 (Age group: 10-12 and 6-8 months) x 2 (Habituating stimuli: [na] vs. [ŋa]) x 2 (Trial type: same vs. change) repeated-measures ANOVA. The analysis revealed no significant effects of Age group, Habituating stimuli or Trial type [Mean$_{SAME}$=5.04s, Mean$_{CHANGE}$=4.66s]. Across both age groups, 11 of the 32 infants tested looked longer to change trials than same trials. The lack of a significant difference in LT to same vs. change test trials indicates that infants across the ages of 6-8

and 10-12 months were unable to discriminate the non-native contrast. The apparent decrease in LT from same to change for the older group was not significant.

4.3.5 Discussion

The results of Experiment 4 are inconsistent with the reorganization hypothesis, which predicts that infants discriminate the non-native contrast at 6-8 months but not at 10-12 months. Previous studies have shown that infants under 10-12 months are capable of discriminating non-native consonant contrasts, which adult listeners have difficulty discriminating, but English-hearing 6-8 month olds (and, as expected, 10-12 month olds) did not reliably discriminate the [na]-[ŋa] contrast. That both groups of infants do not reliably discriminate the [na]-[ŋa] contrast is consistent with the acoustic-perceptual salience hypothesis. This interpretation requires further examination.

Given that the development of the perception of *nasal* contrasts has not been reported in the literature, a possible interpretation of the results of Experiments 3 and 4 is that the development of the perception of nasal consonants is on a different time course than that of oral consonants. That is, perhaps reorganization, where infants' perception develops from discriminating language-general contrasts to language-specific contrasts occurs earlier for nasals. This line of reasoning is further motivated by the well-known finding in the developmental literature that infants' perception of vowels reflects native categories by 6-8 months. Infants at 4 months are able to discriminate non-native vowel contrasts which they no longer discriminate at 6-8 months (Polka and Werker, 1994). Nasals share acoustic properties with vowels (such as sonority) that are different from those of oral consonants. Such a similarity may hint at an earlier time course for reorganization.

This possibility is explored in the next experiment, where English-hearing infants between the ages of 4 and 5 months are presented with native and non-native nasal contrasts.

4.4 Experiment 5: Native and non-native nasal contrast perception at 4-5 months of age

4.4.1 Introduction

In Experiments 3 and 4, if nasal contrasts are perceived like other consonants in infancy, we would have expected reorganization, or the canonical developmental pattern of discrimination with 6-8 month old infants discriminating the non-native contrast that 10-12 month old infants no longer discriminate. That the younger infants failed to discriminate the [na]-[ŋa] contrast suggests two possibilities. The first is that nasal contrasts are subject to reorganization, but the effect occurs earlier than for oral consonant contrasts. The second interpretation appeals to the acoustic-perceptual salience effect introduced in Chapter 1 and shown to be possibly influential in Filipino adults' processing of the native [na]-[ŋa] contrast in Experiment 2.

Experiment 5 examines the first possibility by testing 4-5 month old English-hearing infants on the native [ma]-[na] and non-native [na]-[ŋa] contrasts. If earlier reorganization applies to nasal place discrimination, infants at this very young age group are expected to show the non-native effect, with successful discrimination of both the native [ma]-[na] and non-native [na]-[ŋa] contrasts.

4.4.2   Methods

The procedure and setup were identical to those in Experiments 3 and 4. The stimuli used in the present experiment are the same [ma], [na] and [ŋa] tokens used in Experiments 2, 3 and 4.

4.4.3   Participants

Data from 32 infants were analyzed in the study, 16 infants in the [ma]-[na] group and 16 in the [na]-[ŋa] group. All of the infants were between the ages of 4 months, 0 days and 5 months, 10 days at the time of the study (mean=147 days, range=135 to 160 days). An additional 24 infants participated in the study, but had to be excluded from analysis due to crying or fussiness (n=13), parental or caregiver interference (n=2), experimental error (n=1), habituation in 6 trials (n=1), repeated test trials (n=3), and failure to habituate (n=4). As in Experiments 3 and 4, only infants who received at least 90% Canadian English input according to parental reports were recruited for participation.

4.4.4   Results

In order to assess whether infants attended to the visual stimulus across the duration of the task, a 2 (Contrast: [ma]-[na] vs. [na]-[ŋa]) x 2 (Test: pretest vs. posttest) repeated-measures ANOVA was conducted. There was no effect of either Contrast or Test suggesting that infants maintained attention throughout the task.

Figure 4.4 plots infants' LTs to same vs. change in the two contrast conditions.

English-hearing infants, 4-5 months: [ma]-[na] and [na]-[ŋa]
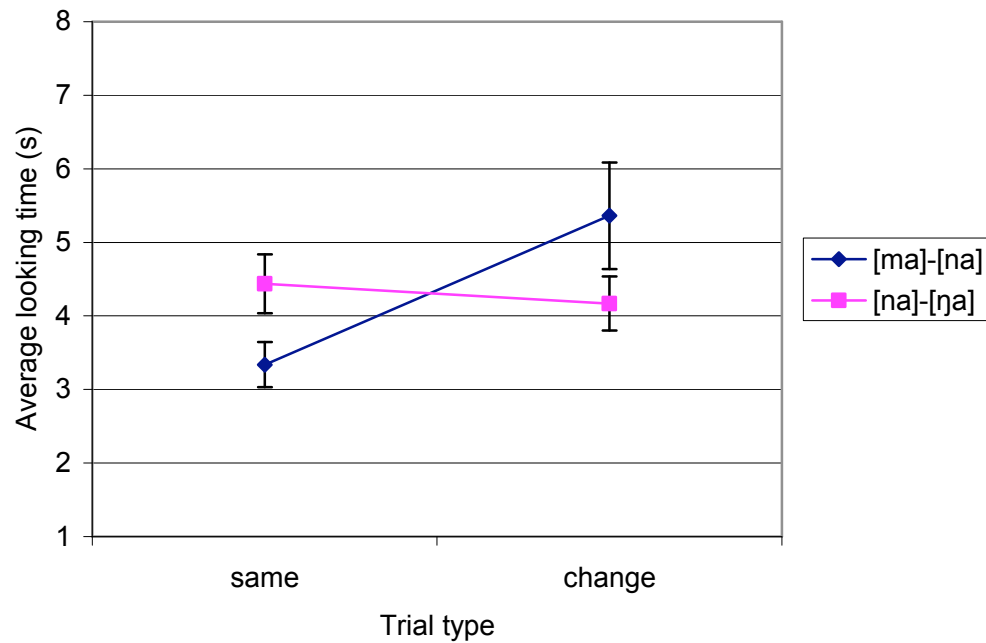
Figure 4.4. English-hearing infants' (at 4-5 months) looking time to same vs. change trials for the native [ma]-[na] and non-native [na]-[ŋa] contrasts.

Because [na] tokens were used as habituating stimuli for both groups ([ma]-[na] and

[na]-[ŋa]), infants were coded with a 4-level between-subjects factor (Habituation-

Contrast: within the [ma]-[na] contrast group, habituation with [ma]; within the [ma]-[na]

contrast group, habituation with [na]; within the [na]-[ŋa] contrast group, habituation

with [na]; within the [na]-[ŋa] contrast group, habituation with [ŋa]). Discrimination of

the two contrasts was examined in a 4 (Habituation-Contrast) x 2 (Trial type: same vs.

change) ANOVA. The analysis showed a main effect of Trial type [$F(1, 28)=6.41$,

$p<0.05$], with infants looking significantly longer to change trials than same trials. There

was also a significant interaction between Trial type and Habituation-Contrast

[$F(1,28)=3.98$, $p<0.05$]. This interaction was further probed with two post-hoc 2 (Trial

type) x 2 (Habituating stimuli) repeated-measures ANOVAs. In the [ma]-[na] contrast condition there was a significant effect of Trial type [$F(1,14)=12.20$, $p<0.005$] with infants looking longer to change trials than same trials [Mean$_{SAME}$=3.34s, Mean$_{CHANGE}$=5.36s]. Within the [ma]-[na] contrast, 13 of the 16 infants tested looked longer to change trials than same trials. There was no significant interaction between Trial type and Habituating stimuli. Because there was no effect of Habituating stimuli, the interaction between Trial type and Habituation-Contrast is driven by the effect of contrast on LT to same and change. This is evident in the figure above.

In the non-native [na]-[ŋa] contrast condition, there was no effect of Trial type nor an interaction between Trial type and Habituating stimuli [Mean$_{SAME}$=4.44s, Mean$_{CHANGE}$=4.17s]. Six of the 16 infants in the [na]-[ŋa] group looked longer to change trials than same trials. The interaction in the main analysis is driven by infants' discrimination of the native contrast.

4.4.5   Discussion

The pattern of nasal consonant discrimination exhibited by 4-5 month old infants is similar to the patterns shown by 6-8 and 10-12 month old infants. Under the reorganization hypothesis infants in this very young age group were predicted to discriminate *both* the native and non-native contrasts. The acoustic-perceptual salience hypothesis predicted that infants would discriminate the acoustically salient [ma]-[na] contrast more accurately than the less salient [na]-[ŋa] contrast. That even young infants did not discriminate the non-native [na]-[ŋa] contrast suggests that reorganization, as observed in the perception of consonant contrasts such as the Hindi dental-retroflex contrast, does not affect the pattern of nasal perception observed in infancy. Rather, the

discrimination pattern shown by English-hearing infants in Experiments 3, 4, and 5 are suggestive of an effect of the acoustic salience of the [na]-[ŋa] contrast.

4.5   General discussion

Taken together, the results of Experiments 3, 4 and 5 paint a picture of [na]-[ŋa] perception different from the reorganization effect in the development of other consonant perception. Under the reorganization hypothesis, infants were expected to show the now well-established perceptual reorganization that occurs for speech contrasts in the first year of life, with successful discrimination of the non-native contrast at some age in young infancy. That English-hearing infants, across their first year, successfully discriminated the native [ma]-[na] contrast, suggests that the lack of discrimination of [na]-[ŋa] was not due to a general inability to discriminate nasal place of articulation.

The acoustic-perceptual salience account introduced in Chapter 1 and suggested by the performance of Filipino-speaking adults in Experiment 3 could potentially explain the current set of results from English-hearing infants. However, the appeal to an acoustic-perceptual salience effect can only be justified if the reorganization hypothesis is ruled out in a cross-language design. That is, only if discrimination of the [na]-[ŋa] contrast proves difficult for young Filipino-hearing infants can we appeal to the acoustic-perceptual salience hypothesis. In Chapter 5, I present the cross-language analog to the English-infant studies (Experiments 3, 4 and 5) by testing Filipino-hearing infants on the *native* [na]-[ŋa] contrast.

CHAPTER 5: Perception of the [na]-[ŋa] contrast by Filipino-hearing infants

5.1   Introduction

The previous chapters have shown that English-hearing infants at 10-12, 6-8 and 4-5 months do not reliably discriminate the non-native [na]-[ŋa] contrast. English-hearing infants in these age ranges were also shown to successfully discriminate the native [ma]-[na] contrast. These results stand in contrast to the reorganization pattern observed in the development of oral stop and vowel perception. Rather, I interpret the results as suggestive of an acoustic-perceptual salience account for nasal place perception, whereby perceptually less salient nasal place contrasts are more difficult to discriminate in infancy than nasal place contrasts that show a more robust acoustic separation in the perceptually relevant F2xF3 space.

The results of Experiments 4 and 5 beg the question of whether *any* group of infants reliably discriminate the [na]-[ŋa] contrast. Previous research overwhelmingly suggests that infants home in on the phonologically relevant consonant contrasts from their ambient language by the end of their first year. In line with this view, the present

chapter presents an experiment assessing the discrimination of the [na]-[ŋa] contrast by infants for whom the contrast is *native*. In Experiment 6 in the present chapter, I test Filipino-hearing infants at 10-12 and 6-8 months of age on their discrimination of the native [na]-[ŋa] contrast. The acoustic-perceptual salience hypothesis predicts that Filipino-hearing infants will not discriminate the [na]-[ŋa] contrast as well as English-hearing infants discriminated the [ma]-[na] contrast.

According to the reorganization pattern, Filipino-hearing infants are predicted to discriminate the contrast at both ages (6-8 and 10-12 months) for the contrast is native and productive in the infants' ambient language. If infants show an effect of the acoustic-perceptual salience of the [na]-[ŋa] contrast they may not show reliable discrimination. The acoustic-perceptual salience account makes no predictions regarding the time course of infants' ability to discriminate the [na]-[ŋa] contrast. There is evidence, however, that listeners may initially show difficulty in discriminating the contrast in infancy and only later in adulthood show reliable discrimination. Recall that Polka, *et al.* (2001) showed that experience with the native language facilitated [d]-[ð] discrimination in English speakers; adult listeners discriminated the contrast that was difficult for English-hearing infants between 6-12 months of age.

Previewing the results, the findings of Experiment 6 suggest that native-language experience allows the Filipino-hearing infants to disentangle the [na]-[ŋa] contrast only in later infancy (10-12 months). Filipino-hearing infants do not reliably discriminate the native [na]-[ŋa] contrast at 6-8 months. This result is suggestive of an acoustic-perceptual salience influence on the perception of the [na]-[ŋa] contrast in infancy.

5.2   Methods

5.2.1   Setup, Procedure and Stimuli

The setup and procedure were identical to those in Experiments 3, 4 and 5 (Chapter 4). The stimuli used in the present experiment are the same [na] and [ŋa] tokens used in Experiments 2, 4, and 5.

5.2.2   Participants

Data from 20 Filipino-hearing infants were analyzed in the study (12 10-12 month olds and 10 6-8 month olds[16]). The older group were between the ages of 10 months, 1 day and 12 months, 20 days (mean=339 days, range=302 to 380 days). The younger group were between the ages of 6 months, 4 days and 8 months, 22 days (mean=219 days, range=188 days to 262 days). An additional 11 infants participated in the study, but had to be excluded from analysis due to their failure to habituate (n=1), habituating in 6 trials (n=2), crying/fussiness (n=6), refusing to look at the TV monitor (n=1), and repeated test trials (n=1).

Many Filipinos in the Vancouver area spoke Filipino (Tagalog) as well as other Austronesian languages of the Philippines. For many families Filipino was used as a *lingua franca* within larger Filipino social networks, while an ethnically and regionally specific Austronesian language (like Cebuano, Ilocano, etc.) was used at home. Infants

---

[16] The lower number of Filipino-hearing infants in the study (relative to the English-hearing infant participants) was due to a number of factors, mostly related to the socio-economics of infant participation in scientific study in general. The overwhelming majority of parents/caregivers who have the time and money to bring their child in to the lab for an unpaid study were Anglo-Canadians from the more well-to-do neighborhoods of Vancouver close to UBC. Larger concentrations of Filipinos and other immigrant communities are found further away from the university in East Vancouver. Additionally, many Filipino parents expressed interest in participating but were unable to leave employment obligations.

who were exposed to *any* Austronesian language of the Philippines were allowed to participate in the study[17]. Recall that all of the Austronesian languages of the Philippines show a robust presence of the velar, alveolar and bilabial nasal places in syllable-onset position (see section 1.3.1). Section 1.3.2 gives the relative frequency of nasal place phonemes in standard Filipino. The relative frequency of these phonemes in the different languages of the Philippines is unknown.

In order to assess the amount of infants' Filipino exposure, a brief language-background questionnaire was administered to the accompanying parent/guardian. This questionnaire was similar to those used by infant speech perception labs interested in quantifying the level of bilingualism in a particular household. Only infants who received at least 50% exposure to a Filipino language and whose mother was born and raised in the Philippines and was a native speaker of a Filipino language were allowed to participate.

5.3   Results

Figure 5.1 plots infants' average looking time to same and change trials according to age group. The plot clearly shows that infants in the older group look reliably longer to change than same trials, while infants in the younger group look equally to same and change trials.

---

[17] Restricting the subject pool to infants being exposed to only Filipino (Tagalog) would have resulted in an extraordinarily long experiment completion time as Tagalogs make up a small percentage of Vancouver-area Filipinos.

Filipino infants, 10-12 & 6-8 months
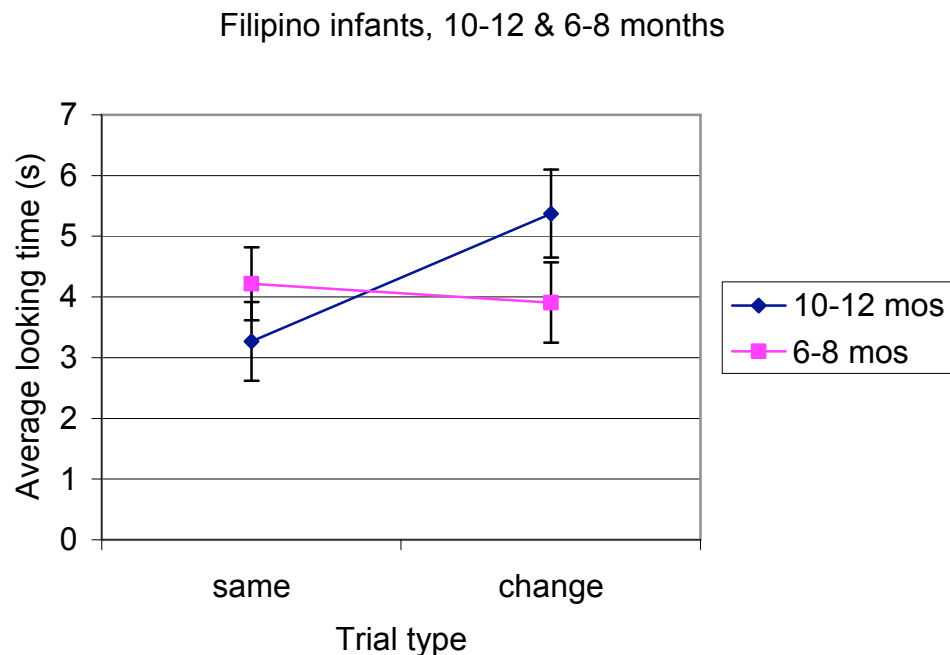
Figure 5.1 Filipino-hearing infants' (at 4-5 months) looking time to same vs. change trials for the native [na]-[ŋa].

In order to assess whether infants attended to the visual stimulus across the duration of the task, a 2 (pretest vs. posttest LT) x 2 (Age group: 10-12 vs. 6-8 months) mixed ANOVA was conducted. There was no effect of either test type or age group, suggesting that infants maintained attention throughout the task.

Discrimination results were analyzed in a 2 (age group: 10-12 vs. 6-8 months) x 2 (habituating stimuli: [na] vs. [ŋa]) x 2 (trial type: same vs. change) ANOVA. The analysis showed a main effect of Trial type [$F(1, 17)=9.84, p<0.01$] with infants looking significantly longer to change trials than same trials. The effect of Trial type can only be interpreted in light its significant interaction with Age group [$F(1,17)=17.44, p<0.005$]. The interaction was further probed with a series of paired samples t-tests. Within the 6-8 month group, there was no significant difference in infants' looking time to same and

change trials [Mean$_{SAME}$=4.21s, Mean$_{CHANGE}$=3.91s]. Four out of the 8 infants in the 6-8 month-old group looked longer to change trials than same trials. There was a significant difference in looking time to same vs. change trials for 10-12 month olds [$t(11)$=-5.71, p<0.001] with infants looking longer to change trials than same trials [Mean$_{SAME}$ =3.27s, Mean$_{CHANGE}$ =5.37s]. Eleven of the 12 infants in the 10-12 month-old group looked longer to change trials than same trials.

5.4   Discussion

Filipino-hearing infants in the present study were shown to discriminate the [na]-[ŋa] contrast only towards the end of the first year of life. Younger infants did not show reliable discrimination. This perceptual development of a native contrast, from poor discrimination in young infancy to successful discrimination in later infancy, is different from the typical pattern described in Chapter 1 in which younger and older infants discriminate native contrasts. Rather, these data suggest that some native consonant contrasts may be difficult for young infants to resolve, with native-language experience being necessary to separate the acoustically difficult contrast into perceptually relevant categories.

Taken together, the results of the present experiment and Experiments 3, 4, and 5, are explained by a combination of language experience and acoustic-perceptual salience for the discrimination of nasal place contrasts in infancy Aspects of these findings were tentatively predicted on the basis of phonological data (Chapter 1), acoustic measures (Chapter 2), and the adult findings from Experiment 2 (Chapter 3). Recall that, in Experiment 2, adult Filipino listeners discriminated the [ma]-[na] contrast better than the [na]-[ŋa] contrast, even though both contrasts are phonemic in Filipino. The next

experimental chapter further tests adults' discrimination and the results provide even

stronger evidence that the [na]-[ŋa] contrast is less salient than the [ma]-[na] contrast,

even for native listeners.

CHAPTER 6: Assessing the perceptual salience of Filipino nasal-onset contrasts

6.1   Introduction

The previous experiments investigated whether perception of native and non-native nasal contrasts develops in a manner similar to the well-known findings for the development of the perception of most other consonant contrasts. The results showed that, unlike previous studies of the development of cross-language speech perception showing non-native consonant contrasts to be discriminable by young infants under 10 months of age, the English-hearing infants in the present set of studies poorly discriminate the non-native [na]-[ŋa] contrast. Additionally, Filipino-hearing infants under 10 months of age, for whom the contrast is *native* show difficulty in discriminating [na] from [ŋa], suggesting that the presence of the contrast in the ambient language does not affect its perception in early infancy. Rather, the explanation for the current pattern of development considers the relative acoustic distinctiveness and resulting perceptual salience of the contrasts as well as the infants' experience with the contrast towards the end of their first year. It is argued that infants under 10 months of age exhibit difficulty

in discrimination precisely because the members of the contrast do not exhibit large acoustic differences. The [ma]-[na] contrast, on the other hand, is acoustically robust (with both murmur and transition cues showing significant differences) and is successfully discriminated by infants as young as 4-5 months old. Infants' whose language does exhibit the contrast only reflect the contrast in their discrimination later in infancy.

The present chapter presents an experiment designed to assess the validity of the interpretation of the infant results by testing the relative perceptual salience of the two nasal contrasts in question. By showing that the two *native* nasal contrasts, [ma]-[na] and [na]-[ŋa], are perceived differently by adult listeners whose phonology reflects the contrasts, we can strengthen the claim that perceptual salience is a factor in the development of the perception of nasal consonants. That is, if the salience argument is to carry the bulk of the explanatory weight for the observed developmental pattern, it should have measurable effects in adult perception as well. A hint of such an effect is found in the results of the initial adult cross-language experiment. Experiment 2 (Chapter 3) showed that, although Filipino listeners successfully discriminate both the native [ma]-[na] and [na]-[ŋa] contrasts at rates above 90% correct, the pattern of their performance was similar to the pattern shown by English listeners in that their discrimination of [na]-[ŋa] was significantly less accurate than that of [ma]-[na]. That Filipino listeners, whose native language exhibits all three nasal places of articulation phonemically, showed this pattern begs further exploration. The hypothesis advanced here for the slightly less accurate performance on the [na]-[ŋa] contrast by adult Filipino listeners is that the contrast is simply acoustically less robust than the [ma]-[na] contrast and consequently

perceptually less salient. The adult English-listener pattern can be attributed to a combination of language experience and acoustic-perceptual salience. English-listeners' perception of the Filipino [na]-[ŋa] contrast is therefore consistent with numerous studies of non-native speech perception (see Strange, 1995 for a full review), but the effect of acoustic-perceptual salience cannot be disentangled from the effect of experience.

How do we empirically justify this rather impressionistic notion of perceptual salience? Fortunately the mature human audition and perception systems are capable of resolving a wide range of acoustic-phonetic contrasts, as evidenced in the excellent (>90% correct) performance of Filipino listeners on [na]-[ŋa] in Experiment 1. What needs to be established is the *relative* strength or salience of the contrasts in question. The current experiment tests relative salience under noisy listening conditions. As listening conditions worsen, there should be asymmetrical effects on the perception of the two contrasts, with the more robust contrast being less vulnerable to interference. All else being equal, if there is more of an effect of the noisy listening condition on [na]-[ŋa] than [ma]-[na], this would support the claim that the [na]-[ŋa] contrast is perceptually less salient than the [ma]-[na] contrast and implicate acoustic-perceptual salience as an influencing factor in the developmental pattern observed in the previous chapters.

As most speech is produced and comprehended in less than ideal listening environments, the introduction of noise to the signal, as in this experiment, might be viewed as mimicking this aspect of naturalistic listening situations. Noise has varying effects on the perception of speech and interacts with phonetic features such as place of articulation and manner (for a fuller review see Benkí, 2003), as well as with phonetic context. In their now-classic study of consonant confusions in masking white noise,

96

Miller and Nicely (1955) found that correct identification of place of articulation is greatly reduced by broadband noise in a closed-set identification task. Their findings showed that listeners' confusions were most often found in place of articulation: as the signal-to-noise ratio decreased, place-of-articulation errors in identification increased. Similar results were found in other studies assessing the effects of noise on consonant identification (Wang & Bilger, 1973; Benki, 2003).

Alwan *et al.* (1999) studied the effects of broadband Gaussian noise on the perception of nasal onsets in English ([m] and [n]). They found that listeners' identification accuracy in /a/ contexts was high (>80%) in very noisy (-10dB SNR) conditions, but poor in /i/ contexts even in 5dB SNR conditions. Alwan *et al.* found that formant transitions into the vowel are perceptually more important for identification in /a/ contexts, while both murmur and formant characteristics seem to be important in /i/ contexts.

While the above linguistic studies assess perceptual salience from degraded cues by way of listener accuracy in an identification task, the present study assesses perceptual salience in a discrimination task. The idea here is that the discrimination of contrast is a gateway into language acquisition. Only when a difference limen between phonetic categories is achieved can the infant bootstrap her way into higher linguistic structure. An analysis of the comparison of contrasts is therefore crucial to our understanding of the developmental patterns of perception. The present study is the first to assess the relative salience of different nasal place contrasts (i.e., [m]-[n] as opposed to [n]-[ŋ]) as a function of noise in the signal.

6.2   Methods

6.2.1 Stimuli

The speech stimuli were the same 12 tokens (four tokens for each place) used in Experiments 2-6. The experiment consisted of three listening conditions: one "clean" or unaltered condition and two "noisy" conditions differing in signal-to-noise ratio. Signal-dependent uncorrelated[18] noise was added offline to the clean stimuli using a specially written script for the MATLAB environment following Schoeder (1968). A product of the signal (s), the *SNR* scaling factor and a randomly generated ±1 were added to every sample of the signal as given in Equation 1:

Equation 1. Formula yielding signal-dependent uncorrelated noise following Schoeder (1968).

$$s + \left(\left(\sqrt{\frac{1}{10^{SNR/10}}} \right) s \right) (\pm 1)$$

Special care was taken in determining the appropriate signal-to-noise ratios which would potentially affect the perception of the two contrasts. Pilot studies with native Filipino listeners showed that performance on [na]-[ŋa] (as measured in proportion correct) slightly decreased at 0 dB SNR (approximately 85% correct), yet did not decrease further at -2dB SNR. Increasing the noise level in the signal to -5dB SNR had the most dramatic effect in the pilot studies and so that level was chosen as the noisiest listening condition used in the experiment.

---

[18] Signal-dependent uncorrelated noise has been shown to have perceptual effects similar to broadband noise (Benki, 2003).

6.2.1.1 Trials

The three listening conditions were assembled into blocks: clean (with no additive noise), 0 dB SNR, and -5 dB SNR. Using the Filipino /ma/, /na/, /ŋa/ stimuli from the earlier experiments, four tokens from each category were paired with four tokens from each different category as well as non-identical tokens from the same category. This assembly yielded 100 pairs (of pair types AB, BA, AA, BB). As "same" pairs never consisted of identical tokens (always A_xA_y, never A_xA_x) this approach yields a stimulus set that is unbalanced with more "different" pairs (64) than "same" pairs (36). In order to balance the set, an additional 28 "same" pairs were added to the list of stimuli, yielding 64 "same" pairs and 64 "different" pairs, or a total of 128 trials. Each trial occurred twice for a total of 256 trials per block (blocked by listening condition: clean, -5dB SNR, -10dB SNR). Consistent with previous adult experiment in this study, the interstimulus interval (ISI) was set at 1500 ms.

6.2.2 Participants

Eleven native speakers of Filipino participated in Experiment 7. Participants were recruited through email solicitation to University of Michigan Filipino student associations. Each participant completed a language background questionnaire in order to assess their language use. Five participants reported extensive knowledge of one other Austronesian language of the Philippines such as Ilocano or Cebuano. Only respondents who were fluent in Filipino and reported that they used it on a regular basis were allowed to participate. All of the participants had spent the majority of their life in Philippines, having lived in the US for an average of 5.2 years. One participant had been in the US for 15 years, but reported visiting the Philippines for at least 6 months out of each year

for the past 5 years. All of the participants were fluent in English and reported that
English was the medium of instruction beginning in late elementary school. None of the
participants knew the purpose of the experiment and all were phonetically naive.   None
of the participants reported any speech or hearing problems.

6.2.3 Procedure

The experiment was conducted in a sound-attenuated room in the Phonetics
Laboratory at the University of Michigan. Participants were seated in front of an Apple
iBook laptop computer from which stimuli were presented and responses recorded using a
specially designed presentation instrument for the PRAAT (v.4.3.31)(Boersma &
Weenink, 2005) software package. Participants listened to stimuli over high-quality
Sennheiser HD 625 headphones at a comfortable volume level. All participants were
tested individually.

It has been suggested that using the listeners' native language for the language of
instruction may help them approach the experimental task in a native listening mode
(Jenkins, 1979; Beddor & Gottfried, 1995). For this reason, before the experiment began
each listener spoke briefly in Filipino with a native Filipino speaker (the Filipino
instructor at the University of Michigan) in Filipino. These conversations lasted
approximately five minutes. As in Experiment 2, written instructions were presented to
participants in Standard Filipino . The instructions as well as on-line computer display
labels of "same" and "different" were given in Filipino in order to maximize the
likelihood that listeners were in a "Filipino mode" of listening. These instructions asked
the listeners to imagine they were in the Philippines participating in a speech experiment
where they would be hearing pairs of speech sounds from Filipino. If the sounds were

from the same category in their native language, they were to click the button "same" on the computer screen; if the sounds were from different categories in their native language, they were to click the "different" button. The experiment was self paced, with trials advancing 1 second after a response was made. Listeners were instructed to respond as quickly as possible in order to ensure timely completion of the experiment.

Every listener was tested with the "clean" (non-noisy) stimuli in the first block. The noisy blocks (0db SNR and -5dB SNR) followed, with their order being counterbalanced across participants. The order of the three listening conditions was not randomized in order to avoid effects of noisy perception on the perception of clean trials. After every 64 trials there was a break in presentation and participants were allowed to take a self-regulated break. After each block the listener was allowed to take a 5 minute break as regulated by the experimenter.

6.3   Results

The results of ten listeners were analyzed. The results of one listener (CY) was excluded from analysis due to a misunderstanding of the instructions. Listeners' performance on the AX experiment is presented in Figure 6.1 as a function of their proportion correct on the two types of *different* trials, [ma]-[na] and [na]-[ŋa], according to listening conditions[19]. These results were submitted to a 2x3x2 ANOVA with the

---

[19] Listeners responses were also converted to a bias-free measure of sensitivity, *d'* (Macmillan & Creelman, 2005). This signal detection theoretic measure takes into account the listener's patterns of hits and false alarms to same (within category) and different (between category) trials in computing sensitivity (see section 3.4.2 for details). The *d'* results for the current data were similar to the proportion correct results, that is, the same general pattern was observed, with sensitivity to the [na]-[ŋa] contrast dropping dramatically with increasing noise with high sensitivity to [ma]-[na] across all three listening conditions.

fixed factors: Contrast ([ma]-[na] and [na]-[ŋa]), Listening condition (clean, 0db, and -5dB SNR) and Block order (0dB block first or -5dB block first).

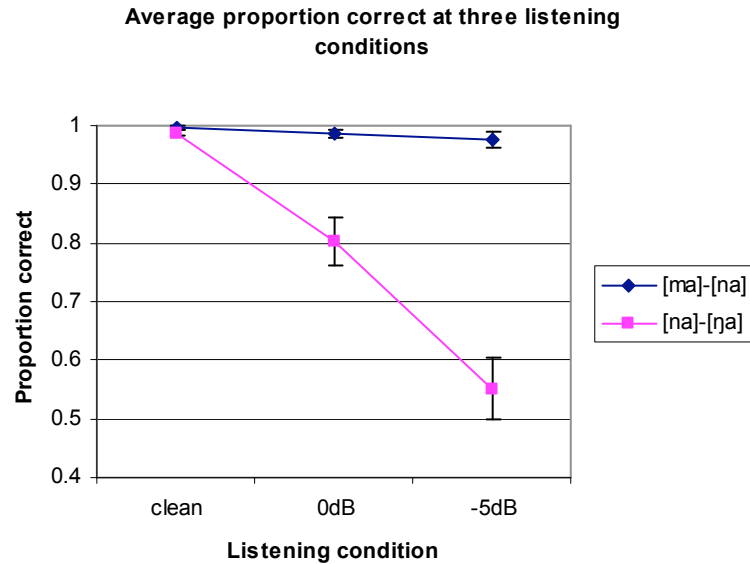**Average proportion correct at three listening conditions**

Figure 6.1. Effects of three listening conditions on the perception of the [ma]-[na] and [na]-[ŋa] contrasts by speakers of Filipino. Error bars represent standard error.

The ANOVA revealed significant main effects of Contrast [$F(1,109)=162.59$, $p<0.001$] and Listening condition [$F(2,109)=58.82$, $p<0.001$]. There was no significant effect of Block order. In addition, there was a significant interaction between Contrast and Listening condition [$F(2,109)=54.62$, $p<0.001$]. These effects were further explored in a series of one-way ANOVAs. Within the clean listening condition there was a small, but nonetheless significant difference between the proportion correct on the two contrasts with listeners performing better on the [ma]-[na] contrast than on the [na]-[ŋa] contrast [$Mean_{[ma]-[na]}=0.9985$, $Mean_{[na]-[ŋa]}=0.9875$; $F(1,39)=6.4$, $p=0.016$]. The effect of contrast was dramatically greater in the 0dB [$Mean_{[ma]-[na]}=0.9953$, $Mean_{[na]-[ŋa]}=0.8031$;

$F(1,38)=38.56$, $p<0.001$] and -5dB [Mean$_{[ma]-[na]}$=0.9906, Mean$_{[na]-[ŋa]}$=0.5513;

$F(1,38)=126.68$, $p<0.001$] listening conditions.

Within the [ma]-[na] contrast, there was no significant effect of Listening condition. Listening condition had a significant effect on the [na]-[ŋa] contrast [$F(2,57)=58.31$; $p<0.001$], with proportion correct dropping from 98% in the clean condition to 80% in the 0dB SNR condition. Accuracy dropped to 55% in the -5dB SNR condition. This effect was investigated in a series of multiple comparisons. Post-hoc (Scheffe) comparisons showed significant differences between clean and 0dB SNR listening conditions ($p<0.001$) and between 0dB and -5dB SNR listening conditions ($p<0.001$).

In general, with the addition of signal-dependent noise, listeners' perception of the [na]-[ŋa] contrast becomes progressively worse. However, this was not the trend for all of the listeners. Figure 6.2 shows individual performance on the [na]-[ŋa] contrast in the three listening conditions.

Effects of noise on discrimination of [na]-[nga]
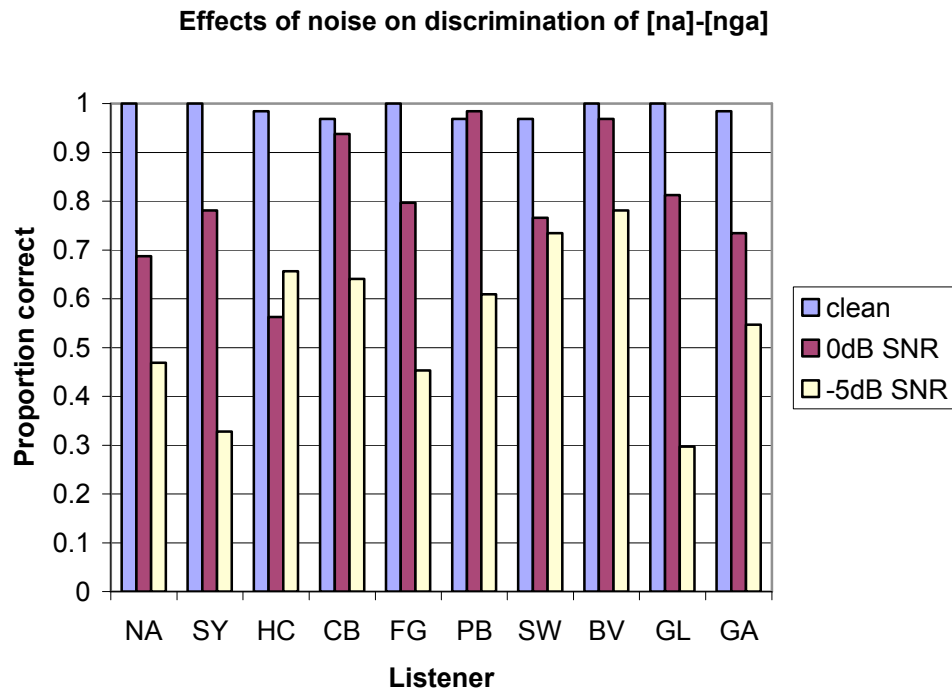
Figure 6.2. Proportion correct on the [na]-[ŋa] at three listening conditions by individual listener.

Notice that equal amounts of noise to signal (0dB SNR) has little effect on [na]-[ŋa] discrimination for listeners CB, PB, and BV. Moreover, for these listeners as well as HC and SW, perception remains above chance (50% correct) level even in the noisiest conditions. Two speakers (SY and GL) performed at substantially below chance levels in the noisiest condition. Below-chance performance is difficult to interpret but suggests that listeners are reasonably confident that different stimuli are actually the same.

6.4   Discussion

The crucial finding from the present experiment is the interaction between contrast and listening condition, with decreases in SNR affecting the perception of the [na]-[ŋa] contrast and *not* the [ma]-[na] contrast. Filipino listeners, for whom both

contrasts are native, performed at near ceiling levels on both contrasts in the clean condition. Consistent with the results of Experiment 2, there was a small but significant difference between listeners' perception of the two contrasts even in the clean condition, although this difference is substantially smaller than in Experiment 2. It is expected that the cross-experiment performance difference is linked to procedural factors. Stimuli in Experiment 1 were presented in a free-field fashion, similar to the infant experiments. It is likely that the closed headphone presentation in the present study is less susceptible to the effects of ambient noise. Additionally, the listeners in the present study adjusted the volume of the presentation to a comfortable level while listeners in Experiment 2 heard free-field trials presented at 68-70 dB SPL.

The more important finding is that listeners' perception of [ma]-[na] remained highly accurate (almost perfect) even in the most adverse listening conditions, while their perception of [na]-[ŋa] degraded to chance performance with increasing noise. Why would the presence of noise in the signal disproportionately affect the [na]-[ŋa] contrast? The most obvious reason is found in the acoustics of these two syllables. Alwan *et al.* (1999) show that in noisy conditions listeners rely on transitional information in categorizing /Na/ syllables (Alwan *et al.*, 1999). Given that [na] and [ŋa] are more similar in the perceptually relevant F2xF3 space (as shown by the discriminant function analysis in section 2.3), we might attribute the source of the perceptual asymmetry in the present study to the acoustic *distance* between the tokens. As with the interpretation of the results of Experiment 2, an interpretation of the results of Experiment 7 must also acknowledge the potential role played by phoneme frequency. That syllable-onset /n/ occurs more frequently than either /m/ or /ŋ/ may also affect adults' perception of the

[ma]-[na] and [na]-[ŋa] contrasts. These effects cannot be teased apart from the present data.

Another possible factor in adult Filipino listeners' enhanced perception of the [ma]-[na] contrast is that all of the listeners in present study were bilingual with the L2 being English. One might argue that listeners in the present study are used to hearing more occurrences of the [ma]-[na] contrast in ambient speech as they all reside in an English-speaking country. This L2 experience might be viewed as being *added* to native language experience, resulting in heightened sensitivity to the more frequent contrast. Although listeners were primed to be attentive to Filipino through the initial conversation with a native Filipino speaker as well as the written experiment instructions, the possible influence of English cannot be discounted.

6.5   Conclusions

The results of the present experiment suggest that the [ma]-[na] contrast is perceptually more salient than the [na]-[ŋa] contrast. This result has implications for the developmental studies presented in the previous chapters and is suggestive of a perceptual salience explanation for certain aspects of patterns of non-native nasal perception in infancy. That the [ma]-[na] contrast is perceptually robust is consistent with typological patterns of nasal consonant distribution in the world's languages, whereby languages which contrast at least two nasal places of articulation do so at the bilabial and dental/alveolar places (Maddieson, 1984). The perceptually less salient [na]-[ŋa] contrast is correspondingly rare in the phonological systems of the world. The interesting question, then, is how does the infant become successful (like their adult counterparts) at discriminating [na] from [ŋa] given that the deck of perceptual salience is

stacked against them? How might the infant resolve the problem of non-prominent

acoustics and weakened perceptual salience in phonetic contrast? The concluding

chapter sketches out potential learning strategies the infant might employ on her journey

to become a competent perceiver of speech.

CHAPTER 7: General discussion and conclusions

7.1   Introduction

The experiments in this dissertation were designed to investigate the developmental perception of nasal consonant contrasts. I approached this topic by offering two distinct hypotheses for predicting infants' discrimination abilities. The *reorganization* hypothesis follows directly from the considerable literature in infant non-native speech perception suggesting that infants' discrimination performance undergoes a reorganization towards the end of their first year as a result of native-language experience. Under this hypothesis, non-native consonant contrasts which adults poorly discriminate are discriminable by 6-8 month olds but are no longer discriminated by10-12 month old learners of that language. The reorganization occurs as the infant tunes her perception to the phonological aspects the ambient (native) language. The reorganization hypothesis, therefore, suggests that perception of nasals will develop in a similar way, with younger infants reliably discriminating both the native and non-native contrasts and older infants successfully discriminating only native nasal contrasts.

The second hypothesis I offered appeals to *acoustic-perceptual salience* as a factor in the development of speech perception in infancy. I argue that speech

perception, regardless of native language experience, reflects low-level acoustic salience resulting in varying degrees of discriminability of speech contrasts. This argument follows in part from the large body of phonetic literature suggesting that acoustics and speech perception help shape phonological patterns. Patterns that are particularly common or particularly uncommon may be due to the inherent discriminability (e.g., due to magnitude of acoustic differences) of the relevant sounds. In the case of nasal place contrasts, alveolar and bilabial nasals are found more often in the world's languages than velar nasals, especially in syllable onset position. Given this asymmetry in distribution, an acoustic-perceptual salience hypothesis would predict that the acoustics and resulting perception of the [na]-[ŋa] contrast are less salient than those of the more common contrast, [ma]-[na]. The prediction for adult discrimination is that listeners' performance on the [ma]-[na] contrast will be significantly better than their performance on [na-[ŋa] for both native and non-native listeners. Correspondingly, the prediction for young infant discrimination under this hypothesis is that the less salient [na]-[ŋa] contrast will be difficult for infants to discriminate regardless of language experience. Language experience is necessary, however, for older infants, whose native language shows a [na]-[ŋa] contrast, to discriminate the acoustically less salient contrast.

In general, the results of the present set of experiments are clearly consistent with an acoustic-perceptual salience hypothesis, in combination with learning due to linguistic experience, for nasal consonant perception. The results of each experiment are described in detail below.

7.2   Summary of results

Chapter 2 revealed that, to the extent that nasal place contrasts are *not* well-differentiated in acoustic space, [na] and [ŋa] are closer in the perceptually relevant F2xF3 space than are either [ma] and [na] or [ma] and [ŋa]. This *acoustic asymmetry*, whereby two phonetic categories are closer together than are other categories, was interpreted as being suggestive of a corresponding perceptual reflex, with the acoustically closer groups being harder to discriminate than the acoustically more "distant" groups.

In **Experiment 2** (Chapter 3) I presented English- and Filipino-speaking adults with the [ma]-[na] and [na]-[ŋa] contrasts in an AX discrimination task. English speakers showed the well-known pattern of non-native speech perception, discussed in section 1.1.1, with near-perfect performance on the native [ma]-[na] contrast and near-chance performance on the non-native [na]-[ŋa] contrast. Filipino speakers overall showed very good performance on both contrasts as /m/, /n/ and /ŋ/ are phonemic in syllable onsets in Filipino. However, the pattern of their relative discrimination of the two contrasts was not different from English speakers' performance; that is, Filipino speakers were significantly better at discriminating [ma]-[na] than [na]-[ŋa]. A likely interpretation of the Filipino listener results is that the [na]-[ŋa] contrast is perceptually less salient than the [ma]-[na] contrast, even though both contrasts are phonemic in Filipino, with the difference in salience following from the acoustic distance between the three places of articulation described in Chapter 2. This interpretation is tempered, however, by the distributional facts for nasal place in spontaneous Filipino speech discussed in section 1.3.2. That /ŋV/ occurs less frequently than either /nV/ or /mV/ suggests that the frequency of these syllables might interact with their perception. These two

110

interpretations, acoustic-perceptual salience and frequency, cannot be disentangled based on the results of Experiment 2. The results of Experiments 1 and 2, combined with the phonological data in Chapter 1, led to the postulation of a working hypothesis of acoustic-perceptual salience as a factor in infant perception of native and non-native nasal contrasts.

In **Experiment 3** (Chapter 4) I presented 10-12 and 6-8 month-old English-hearing infants with the native [ma]-[na] contrast. Both the reorganization and acoustic-perceptual salience hypotheses predict that infants in these two age groups should reliably discriminate the [ma]-[na] contrast. Consistent with both hypotheses, infants at both age groups were successful at discriminating the [ma]-[na] contrast. Infants from the same age groups and language background were then presented with the non-native [na]-[ŋa] contrast in **Experiment 4** (Chapter 4). The reorganization hypothesis predicts that 6-8 month-old infants should discriminate the contrast, while 10-12 month-old infants should fail to make the discrimination. The acoustic-perceptual salience hypothesis makes no definitive prediction regarding the time course of perceptual abilities, but rather suggests that the [na]-[ŋa] contrast is perceptually less salient than [ma]-[na] and this may affect infants' perception even at young ages when infants are expected to discriminate native and non-native contrasts alike. The results of Experiment 4 suggested that, contrary to the reorganization hypothesis, neither the 10-12 month nor the 6-8 month old English-hearing infants reliably discriminated the non-native [na]-[ŋa] contrast. Before appealing to the acoustic-perceptual salience of the contrasts to explain infants' lack of discrimination, even younger English-hearing infants were tested on the native and non-native contrasts. There is evidence in the developmental speech perception literature that

vocalic contrasts become reorganized in the infant's speech perceptual space before consonantal contrasts. Infants have been found to home in on native vowel contrasts by 6-8 months, with the reorganization occurring at 10-12 months for consonants (Polka and Werker, 1994). Given that nasal consonants bear many similarities to vowels, such as high sonority, it seemed reasonable that the reorganization for nasal consonants might occur earlier than for the oral consonants described in the infant speech perception literature. In **Experiment 5** English-hearing infants at 4-5 months were presented the native [ma]-[na] and non-native [na]-[ŋa] contrasts. The reorganization hypothesis would predict that these very young infants would discriminate *both* the native and non-native contrasts, while the acoustic-perceptual salience hypothesis would predict that the more salient contrast, [ma]-[na], will be easier for infants to discriminate than the less salient [na]-[ŋa] contrast. Consistent with the acoustic-perceptual salience interpretation, English-hearing infants at 4-5 months did not reliably discriminate the non-native [na]-[ŋa] contrast, but successfully discriminated the more salient and native [ma]-[na] contrast.

     Experiment 6 (Chapter 5) examined the perception of the [na]-[ŋa] contrast by Filipino-hearing infants, for whom the contrast is native. Again the two working hypotheses make different predictions for infant behavior. The reorganization hypothesis predicts that Filipino-hearing infants at both 10-12 and 6-8 months should successfully discriminate the native contrast because no perceptual reorganization would occur for contrasts that are phonological in the infant's ambient language. As with the other infant experiments, the acoustic-perceptual salience hypothesis would suggest that the [na]-[ŋa] contrast is difficult for infants to resolve even if present in their ambient language. The

results of Experiment 6 suggest that 6-8 month old Filipino infants do not reliably discriminate the native [na]-[ŋa] contrast, while 10-12 month olds successfully resolve the contrast.

The final experiment in this series revisits the adult Filipino-speaker results of Experiment 2 (Chapter 3). Recall that Filipino speakers were successful at discriminating both native contrasts, but their performance of [ma]-[na] was slightly but significantly better than that of [na]-[ŋa]. In **Experiment 7** (Chapter 6) I presented Filipino-speaking adults with the same phonetic stimuli as in Experiment 2 but in three different listening conditions: "clean" (no additive noise), 0dB SNR, and -5dB SNR. The purpose of Experiment 7 was to determine whether the [na]-[ŋa] contrast is indeed perceptually less salient than the [ma]-[na] contrast for adult native listeners. This interpretation would be justified if the degraded listening conditions were to disproportionately affect the discrimination of the [na]-[ŋa] contrast relative to the [ma]-[na] contrast. This is precisely what I found in Experiment 7. Filipino listeners' discrimination of the [na]-[ŋa] contrast fell over 50% with increasing noise in the signal. Their discrimination of the [ma]-[na] contrast remained near ceiling levels across all three listening conditions. These results solidified the claim that the [na]-[ŋa] contrast is perceptually less salient than the [ma]-[na] contrast.

7.3   General discussion

Taken together, the results of the experiments in this dissertation paint a complex picture of adult and infant perception of native and non-native nasal contrasts. The general interpretation of the results of the experiments presented here is that infants' discrimination abilities for syllable-onset nasal place reflect acoustically robust

differences, such as the relatively large acoustic differences between [m] and [n] detailed in Chapter 2. English-hearing infants at all of the tested ages successfully discriminated the acoustically salient [ma]-[na] contrast, but not the less salient [na]-[ŋa] contrast. Further, Filipino-hearing infants discriminated the [na]-[ŋa] contrast at 10-12 months and not at 6-8 months. The Filipino-infants' results suggest that, even when the acoustically less salient contrast ([n]-ŋ]) is present in the ambient language (Filipino), experience is required for the infant to successfully discriminate the contrast. Experience is not necessary for more salient contrasts, as shown by the large literature on the topic. In this study, measures of acoustic salience include size of acoustic differences along perceptually relevant dimensions (Chapter 2) and discriminability for adult native speakers (Experiment 7).

A crucial issue in this study has been the relative roles of native-language experience and acoustic-perceptual salience in the development of nasal consonant perception. More generally, the two hypotheses correspond to two levels of analysis recognized by Aslin and Pisoni (1980) among others as being operative in processing of speech sounds: phonetic or interpretative and sensory and psychoacoustic. The next section further probes the nature of the current results by assessing them in terms of Aslin and Pisoni's (1980) heuristic models of the possible paths infants may take towards mature speech perception.

7.3.1   Induction or attunement?

In their developmental framework, introduced in Chapter 1 (1.4), Aslin and Pisoni (1980) postulate three possible theories that account for the developmental speech perception patterns available at the time: universal theory, attunement theory, and

perceptual learning theory. Which of these developmental theories best describes the

present results from nasal place perception? One might interpret the present data as

demonstrating the "induction" aspect of perceptual learning theory in that infants appear

unable to discriminate the [na]-[ŋa] contrast in early infancy, but experience with Filipino

allows Filipino-hearing infants to make the discrimination by 10-12 months. I contend

that definitively identifying infants' perception of the [na]-[ŋa] contrast as characterizing

induction would be premature. Rather, I claim that the infants in the present study *do*

show evidence for nasal place discrimination, but that it reflects broad nasal place

categories. Recall the plot of [m], [n], [ŋ] tokens in F2xF3 (Figure 2.4) space presented

in Chapter 2. The results of Experiments 3 and 5 (English-hearing infants discrimination

of [ma]-[na] at 10-12, 6-8, and 4-5 months) suggest that infants' initial nasal place

perception is broadly specified, reflecting the categories [m] and [n]. Within Aslin and

Pisoni's (1980) schematic presented in Chapter 1 (Figure 1.1) I would argue that nasal

place perception in infancy is "partially developed" short of specific language

experience; that is, I suggest that infants' perception of [na] and [ŋa] is initially indistinct,

but with specific language experience, separation of the categories emerges. This is

schematized in Figure 7.1. In that figure (a) represents English-hearing infants in the

present study and (b) Filipino-hearing infants. All three nasal place categories are

initially represented in the perceptual system of both groups of infants, but two of these

categories crucially overlap. With appropriate native-language exposure (here Filipino),

reflecting all three categories, the infant separates the initially indiscriminable categories.

In the absence of appropriate native-language exposure, the infant collapses the

previously overlapping categories, thus accounting for English-hearing infants' poor

discrimination of [na]-[ŋa] but successful discrimination of [ma]-[na] at all three age

ranges in the present study. In summary, when the scope of the analysis includes the

class of nasal place contrasts, the current set of results are best supported by Aslin and

Pisoni's (1980) attunement theory. Within this theory, the discrimination of the [na]-[ŋa]

contrast is facilitated in Filipino-hearing infants by the end of their first year.
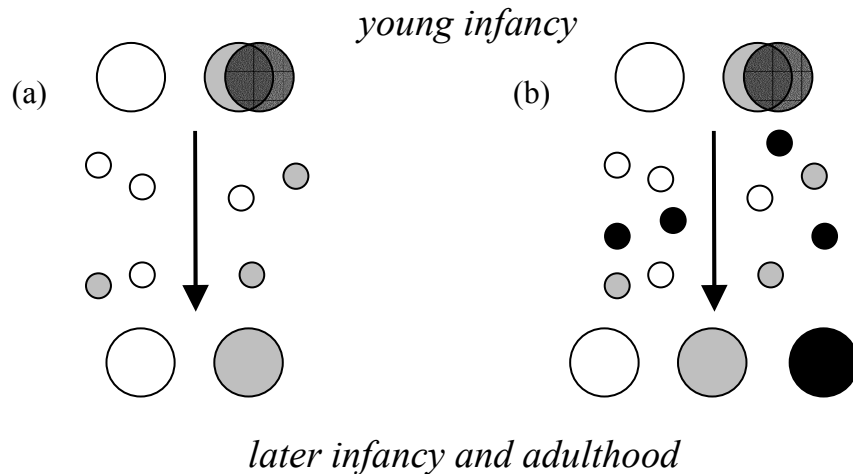

Figure 7.1. Schematic diagram of native and non-native nasal consonant discrimination in infancy.
In younger infancy [na] (grey) and [ŋa] (black) categories remain poorly separated. With specific
language experience (b), the previously undifferentiated category emerges, but not in the absence of
such experience (a).

This interpretation is further reinforced by the frequency of nasals in spontaneous

speech described in section 1.5.2. I showed, based on a corpus of adult-directed

spontaneous speech, that syllable-onset /ŋ/ occurs dramatically less frequently than does

either /n/ or /m/. In fact, syllable-onset /n/ occurs most frequently, accounting for over

50% of syllable-onset nasals in Filipino. Currently, there is no reason to believe that the

distribution of nasal place in infant-directed speech should be different from that in the

adult-directed speech in the corpus study. When we couple the distributional facts of

nasal place input to the infant with the fact that in a perceptually relevant acoustic space

[n] and [ŋ] are similar, it is reasonable to speculate that Filipino infants' *tuning* of their

place categories is evident only later in infancy. The late discrimination of [na]-[ŋa] by

Filipino infants follows from an understanding of infants' category formation as resulting

from a combination of psychoacoustic salience and statistical mechanisms

(Pierrehumbert, 2001; Maye *et al.*, 2002). I have schematized a possible explanation for

the present data in Figure 7.2. As [n] and [m] are the most frequent nasals heard by the

infant, their perceptual space quickly becomes populated by acoustically distant tokens

that are easily discriminable (Time 1). Incoming [ŋ] evidence is sparse and similar to [n],

thus precluding a definitive [ŋ] category distinct from [n] (Time 2). The infant must

receive considerable [ŋ] evidence to form a velar-nasal perceptual category (Time 3).

Given that the [ŋ] input is rare and input [ŋ] has the potential to be treated as [n] by the

perceptual system, an attunement explanation accounts for the present results.

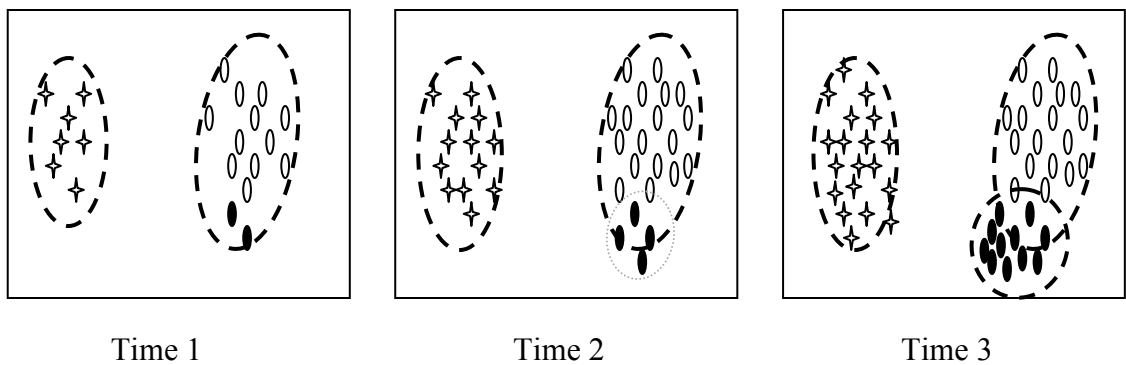
Time 1	Time 2	Time 3

Figure 7.2. Diagram of timecourse of nasal place perception sharpening as the infant receives evidence over time. Stars and open circles represent [m]s and [n]s respectively, with closed circles representing [ŋ]s. Refer to text above for explanation.

The current set of results are in line with recent studies showing the facilitation of

[d]-[ð] and [ɹ]-[l] perception in English speakers (Polka *et al*., 2001 and Kuhl *et al*., 2006

respectively). While Polka *et al*. (2001) showed that [d]-[ð] facilitation occurs after the

first year in English speakers, Kuhl *et al*.'s (2006) results suggest that facilitation occurs

within the first year of life for [ɹ]-[l]. An important distinction between Polka *et al.*'s

(2001) and Kuhl *et al.*'s (2006) results and the results of the present study is the

underlying question being asked. That is, in both Polka *et al.*'s (2001) and Kuhl *et al.*'s

(2006) studies infants were shown to discriminate native and non-native contrasts in

young infancy, with accuracy on the native contrasts increasing and accuracy on the non-

native contrast decreasing with age. The methodology of the present set of infant

experiments does not allow for a direct comparison with either Polka *et al.*'s (2001) or

Kuhl *et al.*'s (2006) studies, both of which employ the Conditioned Head Turn (CHT)

paradigm. Recall that CHT allows the researcher to assess the discrimination ability of

individual infants. With the Visual Habituation (VH) method used in the present set of

experiments, discrimination can only be assessed as a group measure (i.e., we cannot say

if a particular infant discriminates the contrast or not). Additionally, the VH method only

gives the infant one opportunity to show discrimination (one *change* trial compared to a

same trial) while infants in the CHT task are given many more chances. In general, the

VH method is not as sensitive a measure as the CHT and therefore precludes the strong

claim that discrimination of nasal place is induced in infancy. Thus, while the Polka *et

al*. (2001) and Kuhl *et al.* (2006) studies ask whether the initially accurate discrimination

improves, the present study is limited in that we can only assess whether infants, as a

group, discriminate a contrast or not.

7.4   Future directions

The clearest question to emerge from the present set of experiments is how do infants become successful perceivers of *difficult* contrasts. Recent work in developmental speech perception has sought to better understand the mechanism(s) involved in infants' category formation. Research suggests that infants are sensitive to the stochastic distribution of the phonetic variation in the input signal. Maye *et al.* (2002), for example, show that exposing infants to a bimodal distribution of phones from a VOT continuum allows them to discriminate continuum endpoints, which infants exposed to a unimodal distribution do not reliably discriminate. In future work, I would like to explore whether infants' perception of acoustically similar phones, such as /n/ and /ŋ/ falls out from statistical learning mechanism such as that proposed by Maye *et al.* (2002).

A related avenue of inquiry that I would like to pursue is whether conceptual learning aids in discrimination. This idea takes inspiration from classic work in Perceptual Learning (as formulated by Gibson, 1969), work by Waxman and her colleagues on the links between concepts and words in early development (Waxman & Markow, 1995; Balban & Waxman, 1997; Waxman & Booth, 2003), and recent work by Yeung and Werker (2006) attempting to unify conceptual development with infant speech perception. The underlying motivation for an exploration into the links between word learning and speech perception is to uncover whether infants' budding conceptual and word knowledge serves as an imperative for the reanalysis of otherwise difficult acoustic contrasts along perceptually relevant dimensions. Such an imperative may allow

the infant to focus her attention on those aspects of the acoustic signal which were

otherwise ignored.

References

Abramson, A. S., & Lisker, L. (1970). *Discriminability along the voicing continuum: Cross-language tests.* Paper presented at the 6th International Congress of Phonetic Sciences, Prague.

Alwan, A., Lo, J., & Zhu, Q. (1999). *Human and machine recognition of nasal consonants in noise.* Paper presented at the 14th International Congress of Phonetic Sciences, San Francisco.

Anderson, J. L., Morgan, J. L., & White, K. S. (2003). A statistical basis for speech sound discrimination. *Language and Speech, 46*(2-3), 155-182.

Aslin, R. N., & Pisoni, D. B. (1980). Some developmental processes in speech perception. In G. H. Yeni-Komshian, J. F. Kavanagh & C. A. Ferguson (Eds.), *Child Phonology* (Vol. 2: Perception). New York: Academic Press.

Balban, M. T., & Waxman, S. R. (1997). Do Word Labels Facilitate Categorization in 9-Month-Old Infants? *Journal of Experimental Child Psychology, 64*, 3-26.

Bauer, R. S., & Benedict, P. K. (1997). *Modern Cantonese Phonology*. Berlin: Mouton de Gruyter.

Beddor, P. S., & Gottfried, T. L. (1995). Methodological issues in adult cross-language research. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Baltimore: York Press.

Beddor, P. S., Krakow, R. A., & Lindemann, S. (2001). Patterns of perceptual compensation and their phonological consequences. In E. Hume & K. Johnson (Eds.), *The Role of Speech Perception in Phonology*. San Diego: Academic Press.

Beddor, P. S., & Strange, W. (1982). Cross-language study of perception of the oral-nasal

distinction. *Journal of the Acoustical Society of America, 71*, 1551-1561.

Benki, J. R. (2003). Analysis of English nonsense syllable recognition in noise.

Phonetica, 60, 129-157.

Best, C. (Ed.). (1993). *Emergence of language-specific constraints in perception of non-

native speech: A window on early phonological development*. The Netherlands:

Kluwer Academic Publishers.

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*(3), 345-360.

Blandon, R. A., & Lindblom, B. (1981). Modeling the judgment of vowel quality

differences. *Journal of the Acoustical Society of America, 69*, 1414-1422.

Blumstein, S. E., & Stevens, K. N. (1979). Acoustic invariance in speech production:

Evidence from measurements of the spectral characteristics of stop consonants.

Journal of the Acoustical Society of America, 66, 1001-1017.

Boersma, P., & Weenink, D. (2005). Praat: doing phonetics by computer (Version

4.3.31).

Burnham, D. (1986). Developmental loss of speech perception: Exposure to and

experience with a first language. *Applied Psycholinguistics, 7*, 207-239.

Carlson, R., Granström, B., & Klatt, D. H. (1979). *Vowel perception: The relative perceptual salience of selected acoustic manipulations* (No. STL-QPSR 3-4/1979). Stockholm, Sweden: Royal Institute of Technology.

Carrier, J. L. (1979). *The interaction of morphological and phonological rules in Tagalog: a study in the relationship between rule components in grammar.* Unpublished Doctoral dissertation, MIT, Cambridge, MA.

Cho, T., & McQueen, J. (2006). Phonological versus phonetic cues in native and nonnative listening: Korean and Dutch listeners' perception of Dutch and English consonants. *Journal of the Acoustical Society of America, 119*(5), 3085-3096.

Cohen, L. B. (2004). Uses and misuses of habituation and related preference paradigms. *Infant and Child Development, 13*, 349-352.

Cohen, L. B., Atkinson, D. J., & Chaput, H. H. (2004). Habit X: A new program for obtaining and organizing data in infant perception and cognition studies (Version 1.0). Austin: University of Texas.

Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M., & Gesterman, L. J. (1952). Some experiments on the perception of synthetic speech sounds. *Journal of the Acoustical Society of America, 24*(6), 597-606.

Delattre, P. C., Liberman, A. M., & Cooper, F. S. (1955). Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America, 27*(4), 769-773.

Eilers, R. E., & Minifie, F. D. (1975). Fricative discrimination in early infancy. *Journal of Speech and Hearing Research, 18*, 158-167.

Eilers, R. E., Wilson, W. R., & Moore, J. M. (1977). Developmental changes in speech discrimination in infants. *Journal of Speech and Hearing Research, 20*, 766-780.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorito, J. (1971). Speech perception in infants. *Science, 171*, 303-306.

Fant, G. (1960). *Acoustic Theory of Speech Production*. The Hague: Mouton.

Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech, 46*(2-3), 245-264.

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

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments & Computers, 35*, 116-124.

Fujimura, O. (1962). Analysis of nasal consonants. *Journal of the Acoustical Society of America, 34*, 1865-1875.

Gibson, E. (1969). *Principles of Perceptual Learning and Development*. New York: Meredith Corporation.

Gottlieb, G. (1976). The roles of experience in the development of behavior and the nervous system. In G. Gottlieb (Ed.), *Neural and Behavioral Specificity* (Vol. 83). New York: Academic Press.

Hardcastle, W. J. (1973). Some observations on the Tense-Lax distinction in initial stops in Korean. *Journal of Phonetics, 1*, 263-271.

Harnsberger, J. D. (1998). *The perception of non-native nasal contrasts: A cross-linguistic perspective.* Unpublished Dissertation, University of Michigan, Ann Arbor.

Harnsberger, J. D. (2000). A cross-language study of the identification of non-native nasal consonants varying in place of articulation. *Journal of the Acoustical Society of America, 108*(2), 764-783.

Haspelmath, M., Dryer, M. S., Gil, D., & Comrie, B. (Eds.). (2005). *The World Atlas of Language Structures*. Oxford: Oxford University Press.

Hillenbrand, J. (1984). Speech perception by infants: Categorization based on nasal consonant place of articulation. *Journal of the Acoustical Society of America, 75*(5), 1613-1622.

House, A. S. (1957). Analog studies of nasal consonants. *Journal of Speech and Hearing Disorders, 22*, 190-204.

Hunter, M. A., & Ames, E., W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In C. R. Collier & L. Lipsitt (Eds.), *Advances in Infancy Research*. New Jersey: Ablex.

Jenkins, J. J. (1979). Four points to remember: A tetrahedral model of memory experiments. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of Processing in Human Memory*. Hillsdale, NJ: Erlbaum.

Jusczyk, P. W., Murray, J., & Bayly, J. (1979). *The perception of place of articulation in fricatives and stops by infants.* Paper presented at the Society for Research in Child Development, San Francisco.

Kawasaki, H. (1986). Phonetic explanation for phonological universals: The case of
distinctive vowel nasalization. In J. J. Ohala & J. J. Jaeger (Eds.), *Experimental
Phonology*. Orlando: Academic Press.

Kewley-Port, D. (1982). Measurement of formant transitions in naturally produced stop
consonant-vowel syllables. *Journal of the Acoustical Society of America, 72*(2),
379-389.

Krull, D. (1990). Relating acoustic properties to perceptual responses: A study of
Swedish voiced stops. *Journal of the Acoustical Society of America, 88*(6), 2557-
2570.

Kuehn, D. P., & Moll, K. (1976). A cineradiographic study of VC and CV articulatory
velocities. *Journal of Phonetics, 4*, 303-320.

Kuhl, P. K. (1979). Speech perception in early infancy: Perceptual constancy for
spectrally dissimilar vowel categories. *Journal of the Acoustical Society of
America, 66*, 1668-1679.

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), 13-21.

Kurowski, K. M., & Blumstein, S. E. (1987). Acoustic properties for place of articulation
in nasal consonants. *Journal of the Acoustical Society of America, 81*(6), 1917-
1927.

Kurowski, K. M., & Blumstein, S. E. (1993). Acoustic properties for the perception of
nasal consonants. In M. K. Huffman & R. A. Krakow (Eds.), *Phonetics and*

Phonology: Nasals, Nasalization, and the Velum (Vol. 5, pp. 197-222). San Diego: Academic Press.

Larkey, L. S., Wald, J., & Strange, W. (1978). Perception of synthetic nasal consonants in initial and final syllable position. *Perception & Psychophysics, 23*(4), 299-312.

Lasky, R. E., Syrdal-Lasky, A., & Klein, R. E. (1975). VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology, 20*(215-225).

Liberman, A. M., Delattre, P. C., Cooper, F. S., & Gerstman, L. J. (1954). The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs, 68*(379).

Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In W. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling*. Dordrecht: Kluwer.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384-482.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A user's guide* (Second ed.). New Jersey: Lawrence Erlbaum.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Malécot, A. (1956). Acoustic cues for nasal consonant. An experimental study involving a tape-splicing technique. *Language, 32*(2), 274-284.

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

Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America, 27*, 338-352.

Miyawaki, K., Strange, W., Verbrugge, R. R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics, 18*, 331-340.

Narayan, C. R. (2004). A phonological asymmetry and its perceptual consequence: The case of the velar-nasal onset. *Journal of the Acoustical Society of America, 116*(4), 2572.

Newman, R., Sawusch, J., & Luce, P. (1999). Underspecification and phoneme frequency in speech perception. In M. Broe & J. B. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V: Language Acquisition and the Lexicon*. Cambridge, UK: Cambridge University Press.

Nittrouer, S., & Miller, M. E. (1997). Predicting developmental shifts in perceptual weighting schemes. *Journal of the Acoustical Society of America, 101*(4), 2253-2266.

Nord, L., & Sventelius, E. (1979). *Analysis and prediction of difference limen data for formant frequencies* (No. STL/QPSR 3-4/1979). Stockholm, Sweden: Royal Institute of Technology.

Ohala, J. J. (1981). *The listener as the source of sound change.* Paper presented at the Parasession on Language and Behavior, Chicago.

Panganiban, J. V. (1972). *Diksynaryo-Tesauro Pilipino-Ingles*. Quezon City, Philippines: Manlapaz Company.

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*, 325-345.

Pierrehumbert, J. B. (2001). Stochastic phonology. *Glot International, 5*(6), 195-207.

Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics, 13*(2), 253-260.

Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d/-/edh/ perception: Evidence for a new developmental pattern. *Journal of the Acoustical Society of America, 109*(5), 2190-2201.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20*(2), 421-435.

Qi, Y., & Fox, R. A. (1992). Analysis of nasal consonants using perceptual linear prediction. *Journal of the Acoustical Society of America, 91*(3), 1718-1726.

Recasens, D. (1983). Place cues for nasal consonants with special reference to Catalan. *Journal of the Acoustical Society of America, 73*, 1346-1353.

Repp, B. H. (1986). Perception of the [m]-[n] distinction in CV syllables. *Journal of the Acoustical Society of America, 79*(6), 1987-1999.

Rubino, C. R. G. (2000). *Ilocano dictionary and grammar : Ilocano-English, English-Ilocano*. Honolulu: University of Hawai'i Press.

Schoeder, M. R. (1968). Reference signal for signal quality studies. *Journal of the Acoustical Society of America, 44*, 1735-1736.

Seitz, P., McCormick, M., Watson, I., & Blandon, R. A. (1990). Relational spectral features for place of articulation in nasal consonants. *Journal of the Acoustical Society of America, 87*, 351-358.

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics, 17*, 3-45.

Stevens, K. N. (1998). *Acoustic Phonetics*. Cambridge, MA: MIT.

Stevens, K. N., & Keyser, S. J. (1989). Primary feature and their enhancement in consonants. *Language, 65*(1), 81-106.

Strange, W. (Ed.). (1995). *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Baltimore: York Press.

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

Swets, J. A. (1996). *Signal detection theory and ROC analysis in psychology and diagnostics: collected papers*. Mahwah, N.J.: Lawrence Erlbaum Associates.

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

Tsushima, T., Takizawa, O., Sasaki, M., Shiraki, S., Nishi, K., Kohno, M., et al. (1994). *Discrimination of English /r-l/ and /w-y/ by Japanese infants at 6–12 months: Language-specific developmental changes in speech perception abilities*. Paper presented at the International Conference on Spoken Language Processing.

Wang, M. D., & Bilger, R. C. (1973). Consonant confusions in noise: a study of perceptual features. *Journal of the Acoustical Society of America, 54*, 12481266.

Waxman, S. R., & Booth, A. E. (2003). The origins and evolution of links between word
learning and conceptual organization: New evidence from 11-month-olds.
Developmental Science, 6(2), 130-137.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories:
evidence from 12- to 13-month-old infants. *Cognitive Psychology, 29*, 257-302.

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998).
Acquisition of word-object associations by 14-month-old infants. *Developmental
Psychology, 34*(6), 1289-1309.

Werker, J. F., Gilbert, J. H. V., Humphrey, K., & Tees, R. C. (1981). Developmental
aspects of cross-language speech perception. *Child Development, 52*, 349-353.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial
capabilities and developmental change. *Developmental Psychology, 24*(5), 672-
683.

Werker, J. F., Shi, R., Desjardins, R., Pegg, J. E., Polka, L., & Patterson, M. (1998).
Three methods for testing infant speech perception. In A. M. Slater (Ed.),
Perceptual development: Visual, auditory, and speech perception in infancy.
London: UCL Press.

Werker, J. F., & Tees, R. C. (1983). Developmental change across childhood in the
perception of non-native speech sounds. *Canadian Journal of Psychology, 37*,
278-286.

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

Werker, J. F., & Tees, R. C. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America, 75*(6), 1866-1878.

Yeung, H. H., & Werker, J. F. (2006). *Object-based phonetic categorization of speech: a mechanism for functional reorganization.* Paper presented at the International Conference on Infancy Studies, Kyoto.

Zuraw, K. R. (2000). *Patterned exceptions in phonology.* Unpublished Dissertation, University of California, Los Angeles.