Is perceptual reorganization affected by statistical learning?

Dutch infants' sensitivity to lexical tone

Liquan Liu / 3301249
UiL-OTS / Utrecht University
Master Thesis for Linguistics: The Study of the Language Faculty

Supervisor: René Kager
Independent reader: Desiree Capel
July 2010

Dedicated to my grandfather in heaven and grandmother on earth
I love you wherever whenever whatever.

Abstract

At least two things are known with respect to infants' phonemic acquisition: they can track distributional information from the ambient speech input (Maye et al., 2002; Maye et al., 2008), and they experience perceptual reorganization (PR) after which their sensitivity towards non-native speech contrasts greatly decreases (Werker & Tees, 2002). However, it is largely unknown how PR may affect the ability to track sound distributions.

Few previous studies have addressed the statistical learning of lexical tones in infants. It remains unclear whether non-tone-language-learning infants can discriminate tonal contrasts if only provided with some 'right' type of distributional input. Also, although previous studies suggest that tonal PR occurs between 6 and 9 months of age (Mattock & Burnham, 2006; Mattock et al., 2008), the range of tonal contrasts tested to support this suggestion is fairly small. It is likely infants acquire native phonemic inventory at different points in time due to frequency of exposure and the psycho-acoustic difficulty of individual phonemic categories. Hence, multiple tonal contrasts should be studied before concluding on a time window for tonal PR. Moreover, no previous study has combined statistical learning with PR in infants. To solve the issues above, the research questions of this study are: 1) Does statistical learning facilitate infants' discrimination of a non-native tonal contrast? 2) If yes, is this ability affected by tonal PR? 3) When does tonal PR take place?

Ninety-six Dutch infants aged at 5, 11 and 14 months were tested on their perception of a non-native tonal contrast /ta1/–/ta4/ (high-level vs. high-falling) in Mandarin Chinese. A continuum of 8 steps was created for this contrast and two conditions (uni/bimodal) were set up differing solely in the frequency distribution of the stimuli along the continuum. A bimodal distribution is known to facilitate discrimination while a unimodal distribution leads to a decrease in sensitivity. After a 3-minute familiarization phase on one of the two distributions, infants went through a habituation–dishabituation procedure in which their discrimination abilities to the test stimuli were examined.

Results reveal an interesting pattern across ages and conditions. Repeated measures ANOVA shows the Dutch 5-month-old group can distinguish the Chinese tones regardless of the distribution they were exposed to ($p = .020$), while at 11 months only infants trained on the bimodal condition show discrimination ($p = .039$). By 14 months, infants trained on the bimodal condition can no longer distinguish the tonal contrast. The current study supports Mattock and his colleagues' findings about the onset of tonal PR, yet with a slightly different offset since PR may take place later than 9 months in the current study. Future work will address the degree of influence by statistical learning and precise offset of tonal PR. Furthermore, the present study reveals not only the plasticity of PR but also the limitations of statistical learning: to answer the research questions, statistical learning does influence infants' discrimination of the non-native tonal contrast; whereas this influence seems to be reduced before the onset and possibly after the offset of tonal PR.

Contents

Acknowledgement

Half of the acknowledgement goes to lovely people working at Uil-OTS, and the other half goes to warm parents and infants who are willing to contribute to science and join in this research. The first concept of the research topic came up in the course “Language, Speech, Brain” taught by Ivana Brasileiro Reis, and was immediately supported by Desiree Capel with her knowledge and experience in previous M.A. research. Xiaoli Dong and Hugo Quene helped me with the recording and manipulation of the tonal stimuli respectively. Xiaoli Dong, Haifan Lan, Ao Chen, and Wei Wang helped me with the stimuli evaluation. All the members of the babylab group, including Frank Wijnen, Annemarie Kerkhoff, Elise de Bree, Desiree Capel and Josje Verhagen, generously offered suggestions and experiences for the infant experiment. Maartje de Klerk, Eline van Baal and Jorinde Timmer guided me to the babylab facilities and became my hands and voices to help me reach the lovely Dutch parents in Utrecht. Theo Veenker and Warmolt Tonckens helped me set up the testing software and maintain the hardware of the babylab. With an excellent logical mind, Tom Lentz helped me in the reasoning of the adult experiment (Appendix I). I would also like to thank everyone in the phonology group who continuously helped me regardless of the precious time they spent on the topic: Rene Kager, Ao Chen, Diana Apoussidou, Frans Adriaans, Natalie Boll- Avetisyan and Tom Lentz. Thanks to Wei Wang and Glen Chestnut who helped with my English expression and grammar mistakes, and Bert Schouten who instructed me on my internship report. I would like to thank coworker Haifan Lan again for her companion in the basement, which made the study more at ease, and ALL the coworkers in this master’s program, who not only supported me but also brought me valuable opinions in the master’s thesis seminar hosted by Eric Reuland. I would like to send greetings again to my lifelong friend Glen Chesnut and my coworker Desiree Capel who read through my master’s thesis. Last but not least, my greatest thank you to my supervisor Rene Kager. Without Rene’s support my idea would not come into reality. There are many more to whom I would like to thank, like our nice secretaries. I will simply stop here and I just want you to know that I appreciate my life and my working environment so much. Grandpa and grandma, father and mother, this thesis, as well as my every piece of work are for you and always for you. I love you for ever and ever.

1 Introduction

Language acquisition starts before birth and even does not end after puberty. It involves the acquisition of the phonemes and phonology, the words and word meanings, the semantics and syntax of the native or a foreign language. This thesis looks into infants' acquisition of speech sounds during the first year of life and focuses on the potential interaction between distributional input and perceptual changes. Specifically, this thesis investigates whether Dutch infants' speech perception of non-native linguistic tones is affected by statistical information and/or perceptual reorganization.

1.1 Infant speech perception

Language acquisition starts before birth. Newborn infants prefer speech to non-speech (Vouloumanos & Werker, 2007), prefer the language spoken during pregnancy by their mother to other languages (Mehler, Jusczyk, Lambertz, Halstead, Bertoncini & Amiel-Tison, 1988), prefer their own mothers' voices to other female voices (De Casper & Fifer, 1980) and can discriminate between the ascending and descending pitch contours of a certain language (Nazzi, Floccia & Bertoncini, 1998). All of these results show that infants have an early sensitivity to speech sounds, especially to the prosodies and rhythms of the language.

However, more is needed for the infants to fully acquire the native phonemic inventory. They have to not only extract sound segments from the daily speech stream which contains no clear boundary signals, but also capture the phonetic categories and relate them to phonological categories of their native language (Jusczyk, 1992), irrespective of speakers' variability and coarticulation.

1.1.1. Speech contrasts discrimination

Since phonetic categories can be seen as the psychological correlates of the contrastive aspects of phonemes (Maye, 2000), the studies of speech contrast discrimination are crucial. In the beginning, infants are able to detect virtually all natural language speech contrasts[1], the differences of which are

---

[1] Almost but not all, see Polka et al., 2001 for more information.

often beyond adults' recognition. This sensitivity of newborn infants suggests that some experience-independent factor may play a role in early speech perception. With continuous exposure to the native language, infants' speech perception becomes more and more language-specific and they gradually lose sensitivity to the speech sounds which do not reflect the phonological pattern of their native language. Infants' speech perception then becomes categorical, since differences within phonetic categories of the native language are ignored and speech sounds are sorted into categories (see section 1.2). Some early studies addressing this issue focus on the comparison between infant and adults, and others solely investigate infants' perceptual capacities.

Eimas, Siqueland, Jusczyk & Vigorito (1971) showed that infants had some innate perceptual boundaries. One- and four-month-old infants were tested on native voicing consonant contrasts between syllables [ba] and [pa]. Infants habituated on one syllable type showed significant increased in sucking rate when hearing the stimuli from a different (adult) phonemic category, whereas infants habituated and dishabituated on the stimuli that were within one (adult) phonemic category did not show a significant difference. This suggests that these infants group speech sounds into different perceptual categories. Contrasts on place or manner of articulation revealed the same results, reflecting a wide range of categorical discrimination capacities in infants (Eimas, 1974; 1975ab).

Infants can also detect the ordering of information from the speech input. Miller & Eimas (1979) found that infants are sensitive to the combination of elements at both syllabic and segmental levels. For example, infants habituated on a [ba]–[dæ] sequence showed a significant increase in sucking rate when exposed to the similar sequence [bæ]–[da], in which the vowels are reversed.

The contrasts infants can detect may come from non-native phonemic inventories too. The voiced and voiceless contrast used in Eimas et al. (1971) could be discriminated by 2-month-old Kikuyu infants (Streeter, 1976). Similar results were obtained in Trehub's early study (1976): English infants from 5 to 17 weeks are able to discriminate an oral-nasal contrast [pa] and [pã] from French and Polish, and a strident contrast [za] and [řa] from Czech.

Werker & Tees (1984) found that the sensitivity to non-native contrasts declines in the first year of life. English infants were capable of discriminating between the Hindi voiceless unaspirated retroflex and dental contrast [ța]–[ta] and the Thompson glottalized velar and uvular contrast [k̀i]–[q̀i] at 6–8 months. However, their performances declined at 8–10 months, and at 10–12 months, the sensitivity to these non-native contrasts was lost. In comparison, Thompson and Hindi infants distinguished these native contrasts across all age ranges. This pioneering finding is crucial for later discussion in PR.

1.1.2. Speech prosody perception

Prosody in speech refers to intonation, rhythm, stress, tones, etc. The prosody of a language reflects its patterns in the sound structure. This characteristic differs across languages and infants seem to be aware of prosodic cues very early on. As has been mentioned, new-born infants are sensitive to prosodic features conveyed in their mothers' native language (Mehler et al., 1988).

Two-month-old infants can not only discriminate two phonetically identical syllables differing only in pitch contours (Morse, 1972), but also distinguish bi-syllabic segments with stress on the first or second syllable (Jusczyk & Thompson, 1978). Four-month-old infants prefer the expanded intonation contours in infant-directed speech to adult-directed speech differing mainly in fundamental frequency patterns (Fernald & Kuhl, 1987). Eight-month-old infants discriminate simple English sentences differing only in intonation (Kaplan & Kaplan, 1971).

All these findings seem to suggest that infants can extract prosodic features like intonation and stress. This makes the research on speech perception of tone fairly interesting. Although tone is clearly one of the prosodic features, it is quite different from other characters like intonation and stress because it greatly affects the word meanings in a tonal language. Given that word recognition is the ultimate goal for infant speech perception, it seems not unlikely that infants of a tonal language notice this feature at an early age.

1.2 Perceptual reorganization in the first year of life

From discriminating virtually all speech contrasts to distinguishing contrasts within the native phonemic inventory, infants' perceptual sensitivity changes during the second half of the first year of life. This phenomenon is often addressed as perceptual reorganization (PR) to the native language. In other words, PR discusses when infants start to concentrate on the relevant contrasts for the native language and what occurs to the sensitivity to perceive contrasts not present in the native language. The mechanisms underlying PR could be statistical properties of infants' input, phonological categories and perceptual phonetic space, etc (Jusczyk, 1997; Kuhl, Conboy, Padden, Nelson & Pruitt, 2005; Mattock & Burnham, 2006; Pallier, Bosch & Sebastián-Gallés, 1997; Werker & Tees, 2002).

1.2.1 Studies of perceptual reorganization

As has been discussed, Werker & Tees (1984) were the first to show PR in the first year of life. Kuhl and colleagues (Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992) tested 6-month-old American and Swedish infants on two sets of vowel stimuli, the prototype of which represents either English /i/ or Swedish /y/. The findings illustrated that 6-month-old infants had a strong "magnet" effect, namely the perception is prone to the prototype of the phonetic categories of the native language. This did not occur in infants between 4 to 6 months (Polka & Werker, 1994). This indicated that PR for vowels occurred as early as 6 months of age. As for PR of consonants, Pegg & Werker (1997) tested English infants on a [d]–[t] contrast (/d/ in the initial position and /t/ is unaspirated). Results showed that infants of 6–8 months could distinguish the contrast while 10–12-month-olds failed to do so. This finding illustrates that PR occurs in the course of shaping the phoneme inventory of the native language. Werker and colleagues (Werker & Logan, 1985; Werker & Lalonde, 1988) found that infants head their way of building phonological categories by the age of 12 months. FIGURE 1 provides a concise overview of infant speech perception and production in the first year of life.

Besides the decline in sensitivity to non-native phonemic contrasts, PR can also be reflected by how the ambient language environment may facilitate infants' native phonemic perception in the first year of life. Kuhl and colleagues (Kuhl, Stevens, Hayashi, Deguchi, Kiritani & Iverson, 2006) tested

American and Japanese infants on an English /ra/–/la/ contrast existing in English but not in Japanese. Results showed that infants of 6–8 months from both language backgrounds are able to perceive the contrast, while at 10–12 months of age, American infants' ability to discriminate the contrasts increased while this ability is decreased among Japanese infants.

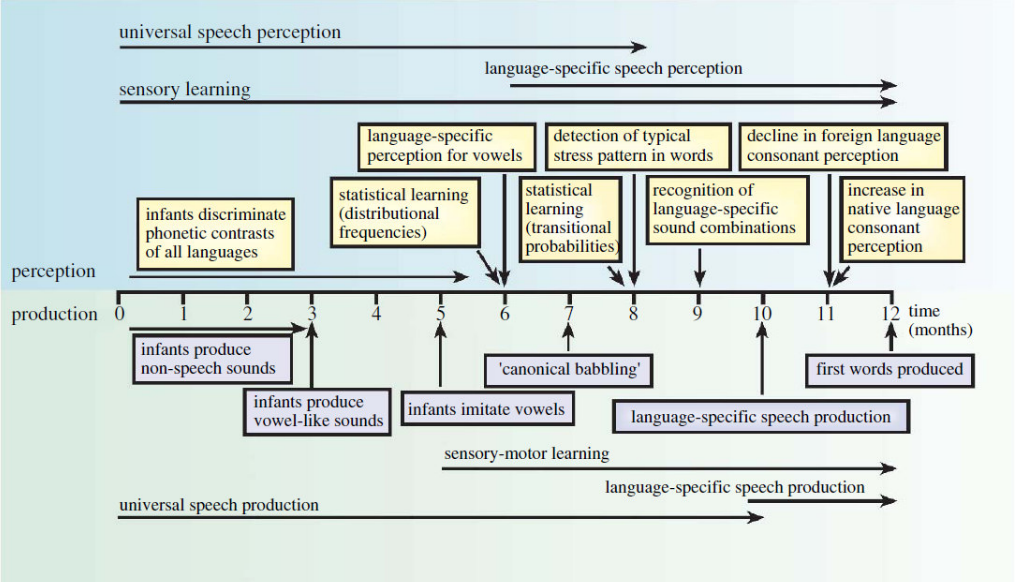


FIGURE 1 Timeline of infant speech perception and production in the first 12 months

Source: Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola & Nelson. (2008)

According to FIGURE 1, PR for different phonetic units varies in time. Generally, PR for native vowels occurs at about 6 months of age while PR of native consonants starts later at around 11 months. However, Werker & Tees (1984) found perceptual change on consonant contrasts at 8–10 months, and their study showed that PR is more likely a gradual turning to the native categories than a drastic change at a certain moment of time. If so, PR of native consonants should be considered to start earlier than 11 months, possibly at 8 months.

An interesting question is why PR of vowels occurs earlier than that of consonants. One possible explanation is that there are usually less vowel categories than consonants in a language, and therefore consonant space is phonetically and perceptually more crowded, requiring more precise perception workload in setting the boundaries and categorizing multiple phonemes, whereas the perceptual space or distance for vowels is wider (Jusczyk, 1997). Another approach could be that vowels are acoustically more salient and easier to discriminate. Note that for new-born infants acoustic changes involving vowel information are more salient than those involving consonants (Bertoncini, Bijeljac-Babic, Jusczyk, Kennedy & Mehler, 1988). A third possibility is that vowels are generally more frequently produced than consonants, resulting in sufficient statistical input at an earlier time. Fourth, it could be due to the functional difference between consonants and vowels (Nespor, Peña & Mehler, 2003; Bonatti, Peña, Nespor & Mehler, 2005). Consonants generally mark lexical contrasts in a language while vowels are initially more helpful in prosodic and phrasal structure learning. Infants' early sensitivity to vowels may be driven by the urge to understand prosody such as phrasing intonation. A fifth explanation would be that infants are sensitive to supra-segmental information carried by vowels (but not consonants), drawing infants' attention in the first place. Finally, it could be that consonantal contrasts are perceived categorically while vowel contrasts are perceived with more gradient. PR might occur relatively easier for gradual perception than for categorical perception, since in the latter case, more information may be needed to form a unimodal-like distribution of certain phonemic category. If this is true, then if lexical tones follow a categorical perception pattern, tonal PR would be later than PR in vowels in terms of time window. Non-speech tones differing in tone onset time were tested in 2-month-old infants by Jusczyk and colleagues (Jusczyk, Pisoni, Walley & Murray, 1980). Results indeed suggest three perceptual categories along the temporal continuum. However, no direct evidence has been obtained for tonal categorical perception in infants so far (Werker & Tees, 2005). All these possibilities may coexist and interact.

PR occurs in phonetic and phonotactic patterns of the native language as well. Jusczyk and colleagues (Jusczyk, Friederici, Wessels, Svenkerud & Jusczyk, 1993) tested 6- and 9-month-old English and Dutch learning infants on a list of carefully selected infrequent words in English and in Dutch. The sounds of these words contained combinations of segments that were either against the phonological rules or not found in the other language. Results showed that at 9 months infants preferred words that reflect the phonotactic patterns of their native language while 6-month-olds did not show this preference. Jusczyk, Luce & Charles-Luce (1994) tested 6- and 9-month old American infants on two sets of non-word lists, one containing segments with frequently occurring phonetic patterns in English and the other with infrequently occurring patterns. Once again infants of 9 months rather than 6 months preferred the lists with high frequency in phonetic patterns of the native language.

It seems that only non-native contrasts that are subject to equivalent classification with a native category are difficult for infants to perceive. Some non-native contrasts may not even have correspondent phonological category in the native language. For instance, a Zulu click consonantal contrast (dental vs. lateral) was well discriminated by English infants of 12–24 months (Best, McRoberts & Sithole, 1988), whereas PR of consonant is supposed to occur at around 11 months of age. Besides that, speech sounds that do not exist in the native phonemic category seem to be difficult to perceive in general across age. As for some contrasts which are acoustically more difficult[2], the PR pattern also differs from the usual assumption. Polka, Colantonio & Sundara (2001) tested 6–8 and 10–12-month-old English and French infants on English /d/–/ð/ stop contrast. Results showed that no single group (not even native English-learning infants) discriminated the contrast well. Considering that English adults can well discriminate the contrast, it is likely that sensitivity of this contrast improves after 12 months with more exposure to the native English language environment.

1.2.2 Accounts and Models related to perceptual reorganization

It has been suggested that the PR may be influenced by the complexity of the phonetic or perceptual space, resulting in a longer process of reorganization (Jusczyk, 1997; Sabourin, Bosch, Sebastián-Gallés & Werker, 2006). A more complex perceptual space may cause more precise perception workload in setting the boundaries and categorizing multiple speech sounds. The PR pattern may also be influenced by phonological status: possible lexical forms require infants to track specific phonological constraints of the native language. Considering that infants start to recognize phonotactic constraints of their native language around 8–9 months of age (Jusczyk et al., 1993), PR may reflect the transformation from phonetic to phonological information and serve as a foundation for the eventual acquisition of the lexicon (Bosch & Sebastián-Gallés, 2005; Werker & Pegg, 1992). Statistical learning may well play a role in PR: the relative frequency of the contrasting phonemes

---

[2] Factors that account for the delay in PR of English /ð/ may involve its acoustic properties, phonotactic uniqueness and lexical knowledge influence on phonetic processing, as Polka et al. (2001) argue.

determines the PR time window of each phonemic category (Anderson, Morgan & White, 2003). Note that both the language-specific phonemic categorization from a crowded phonetic space and phonotactic rule formation may need certain amount of distributional information. Both the frequency account and the perceptual space account go with the finding that PR for vowels occurs earlier than consonants. Last but not least, experience-independent explanations may come into play as one of the factors of PR. An innately guided learning process is likely for the development of speech perception. Moreover, PR can be viewed as a critical/sensitive period (or optimal period, see Werker & Tees, 2005).

A number of theoretical models have been proposed to account for the way speech is perceived. Inevitably, these models need to address infants' speech perception and subsequently the issue of PR. Jusczyk (1993, 1997) proposed the model of "Word Recognition and Phonetic Structure Acquisition" (also known as WRAPSA). Speech input first goes through a preliminary analysis, resulting in the establishment of acoustic properties like spectral and temporal features. Then these properties go through a language-specific weighting scheme acquired from statistical experience of the native language. The weighted signal is further subjected to pattern extraction, a refining process that segments the signal into word-size units, and finally serves as a probe to long-term memory. In this model, PR is seen as selective attention and is explained as stretching and shrinking of distances in psychological space (Jusczyk, 1997). This is caused by the weighting of information.

Kuhl and colleagues (2008) proposed a revised version of the neurologically based speech perception model "Native Language Magnet theory expanded" (NLM-e). In this model PR refers to neurons' commitment to the native language phonetics based on the consistent frequency distribution from the ambient environment. The phonetic exemplars from repeated listening experience define a "prototype" category of phonetic perception. The non-native categories later have to collapse into the committed native ones: the closer they are to the native categories, the harder it is to discriminate them. More evidence should be obtained from a neural-linguistic perspective.

Perceptual Assimilation Model /Articulatory Organ (PAM/AO; Best, McRoberts & Goodell, 2001) is a model based on similarity-detecting mechanisms. It uses gestural phonology, the recovery of vocal tract production, to account for speech perception. The non-native speech sounds are perceived based on how they might be assimilated to the native categories. They are also harder to perceive for infants if the sound contrasts involve the same articulatory organ, as opposed to different articulatory organs. Yet it is difficult to imagine how infants perceive sounds before the babbling phase when they are not capable of producing any sound. So far there is no psychological or neurological evidence showing that infants track the gestural information from birth.

1.2.3 The role of perceptual reorganization in tone acquisition

PR studies on a supra-segmental level are scarce. Jusczyk et al. (1993) conducted an experiment testing 6-month-old American infants on English and Norwegian words which differed in the prosodic organization. A preference was found for the prosody from the native language as early as 6 months. The study by Jusczyk et al. (1994) also revealed that 9-month-old infants preferred the native stress pattern. Unfortunately, fewer studies seem to address the issue of tonal PR.

Mattock & Burnham (2006) tested 6- and 9-month-old Chinese and English infants on non-native lexical tones in Thai as well as musical/non-speech tones. The two contrasts were a rising–low contour–level contrast and a rising–falling contour–contour contrast. The findings pointed to a decline in sensitivity to linguistic but not to non-speech tones in 9-month-old English infants. Mattock, Molnar, Polka & Burnham (2008) further tested 4- and 6-month-old English and French infants on same contrasts and the results showed no decline of sensitivity by that time. These two studies suggest that PR for lexical tone occurs between 6 and 9 months. Intuitively, this onset of tonal PR seems late, given infants' early sensitivity to native language rhythmic structure and intonation.

Although Mattock and colleagues claim that the developmental time-line of PR of tone is around 6 to 9 months, several questions need to be answered. Firstly, in Mattock and his colleagues' studies only two pairs of tonal contrasts were tested. Thus, the range of tonal contrasts tested to support this suggestion is fairly small. Since infants may acquire categories of the native inventory at different points in time due to the psycho-acoustic difficulty and frequency of exposure of the individual categories, multiple tonal contrasts should be studied before concluding on a time window for tonal PR. Secondly, the authors believe that if tone is perceived as a feature of the vowel, the PR time line should be the same for vowel and tone perception. This is however not necessarily true. The view, in which tone is treated as a segmental feature of a vowel, does not hold from a phonological or neurological perspective. Cross-linguistic phonological distributional evidence favours autonomy of tones and vowels. Autosegmental phonology theory (Goldsmith, 1976; 1990) claims different levels of representation of segments and tones. Tone-bearing units can be vowels, but they can also be sonorants. Tones can assimilate between syllables and tonal information can be preserved even when vowels are deleted. In short, tone is more independent and consistent in distributional patterns compared to vowels phonologically. Besides, tones are perceived non-linguistically to non-tonal language speakers from a neurological perspective (Francis, Ciocca, Ma & Fenn, 2008). Last but not least, it has been argued that the mastery of tones occurs earlier than that of segmentals (Zhu & Dodd, 2000; Li & Thompson, 1977). Intuitively, this claim appears to be possible since infants are born with discrimination ability for the prosodic difference between a native (mother's) language and a foreign language with a different rhythmic structure. This shows that the prosodic features are acquired in the prenatal period. The seemingly contradictory observations request further research on a more accurate onset and offset time window of tonal PR, and on how infants perceive each tonal contrast.

No other study has focused on the tonal PR pattern in infants. It remains unclear how tone is acquired and whether or not it is acquired in the same way as other prosodic features. It is intriguing to ask questions such as these: Does early prosodic perception assist in tonal category formation? Does psycho-acoustic saliency influence the acquisition of tones? And does other prosodic information accelerate or impede tone acquisition? The current study may shed light on some, but not all, of these questions.

1.3 Statistical learning in infant phonemic category formation

Statistical learning is sensitive to the absolute and relative frequency of various properties from the input. It may be used in many aspects of language acquisition.

Maye (2000) proposed a distribution-based model to account for the acquisition of native phonemic categories in the first year of life. In this model, phonemic contrasts are acquired by the frequency of contrasting sounds in the speech input in a given phonetic context. Note that this is very close to the account that infants' PR relies on their sensitivity to the statistical properties of the native language (Kuhl, 2000).

Maye, Werker & Gerken (2002) tested 6- and 8-month-old English infants on a consonant contrast — voiced and voiceless unaspirated alveolar stops [t] and [d]. An 8-step continuum was created differing in VOT of the initial consonant from syllable [ta] to [da]. Infants were first divided into two groups, each trained with a type of distribution – unimodal or bimodal, differing in the distributional frequency of stimuli along the continuum. Maye et al.'s goal was to see whether infants were sensitive to the patterns covered by the distribution of speech sounds in a given language. The hypothesis was that infants were able to use statistical information to track the linguistic relevance of the properties of these sounds. In other words, bimodal distribution of sounds would mean that the acoustic property was linguistically important and formed a distinguishable contrast in a language, while unimodal distributions would not be informative for distinguishing some speech sound categories in the language and the variations should be ignored. FIGURE 2 illustrates schematized uni- and bimodal distributions.

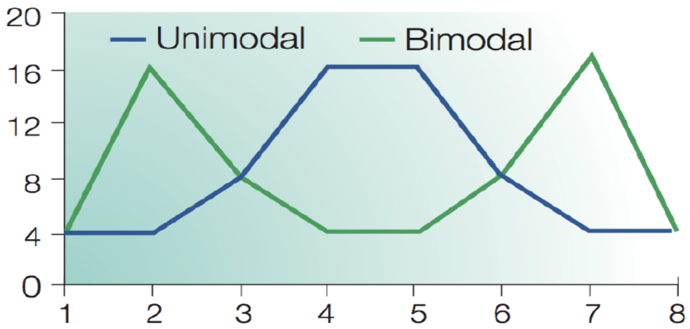


FIGURE 2 Two distributions differ in familiarization frequency (vertical axis) of sounds in a [da]–[ta] continuum (horizontal axis).

Source: Maye et al. (2002)

In the bimodal condition, stimuli on the edges of the continuum (2, 7) are presented more often than stimuli in the middle of the continuum (4, 5) whereas the opposite holds for the unimodal condition. After exposure to either of the two distributions, infants were tested on either alternating (stimuli 1 and 8) trials or non-alternating (stimuli 3 and 6) trials. Note that stimuli 3 and 6 had the same distributional frequency in either modal (8 times in Maye et al., 2002). Results showed that both 6- and 8-month-old infants in the bimodal condition discriminated the test trials while those in the unimodal condition did not perceive the distinction. This means that infants as early as 6 months are sensitive to the statistical distribution of speech sounds in the ambient language which influences speech perception.

Maye, Weiss & Aslin (2008) pushed this idea forward and looked into the enhancement of bimodal condition. Eight-month-old American infants were tested on the discrimination of two Hindi VOT contrasts after training to bimodal distribution, unimodal distribution or non-speech sounds along the 8-step continuum. The design of testing method was revised in this study. After familiarization with one of the two distributions, Maye et al. (2002) used the alternating/non-alternating paradigm to test infants' discrimination ability while Maye et al. (2008) used the habituation and dishabituation paradigm. Results not only showed that bimodal distributional exposure led to a better discrimination of difficult contrasts in the native language (Maye et al., 2002) but also that infants were capable of extracting abstract featural representations from the phonetic input: they successfully distinguished a new contrast that was the featural analogue (e.g. a similar contrast at a different place of articulation) of the trained contrast in the bimodal condition. Table 1 briefly summarizes three situations addressed by Aslin & Pisoni (1980) that reflect the core of Maye et al.'s studies. In short, infants' exposure to regularities within natural languages leads to PR in the form of perceptual decline, maintenance or enhancement of phonemic discrimination (Mattock et al., 2008).

Contrast	Phonetic discrimination		Perception
	6 months	9 months	
Non-native	Yes	No	Decline
Native & Non-native	Yes	Yes	Maintenance
Native	No	Yes	Facilitation

Table 1 Three main types of phenomena under statistical learning

Capel (2008) replicated the study by Maye et al. (2002, 2008) and tested Dutch infants of 10 months on a continuum of Hindi voiced and voiceless retroflex plosive contrast /ɖa/–/ʈa/. The infants were exposed to unimodal or bimodal frequency distribution. Results showed that only Dutch infants who were exposed to the bimodal condition could discriminate the contrast. The interpretation was that infants in the bimodal condition formed two categories in the speech contrast they heard, while infants in the unimodal condition formed only one. These findings were in clear accordance with Maye et al.'s study.

In a recent paper, Yoshida, Pons, Maye & Werker (2010) tested 10-month-old English-learning infants' discrimination to uni/bimodal distributional information on a voicing (a /da/–/ta/ continuum) or place-of-articulation (a synthetic Hindi retroflex–dental continuum) speech sound distributions. The same design and paradigm was used as in the study of Maye et al. (2002). Results showed that infants at 10 months lost their sensitivity to the non-native speech sound distributions. However, when the familiarization time of place-of-articulation distinction was doubled (longer exposure), infants were able to discriminate the sounds. The authors argued that distributional learning at 10 months of age remained effective, but more difficult than before PR (6–8 months) and they also proposed that distributional learning may be an underlying mechanism of PR.

Statistical learning seems quite plausible to account for the development of speech perception at an early age. Going back to FIGURE 1, it can be observed that statistical learning goes with PR quite well – it co-occurs with language-specific perception. The potential enhancement and impedance of acquisition based on distributional input may also explain why sometimes PR does not follow the regular pattern which occurs at around 10 to 11 months of age for consonants. For instance, the English /d/–/ð/contrast appears to be quite difficult for English infants to perceive initially. Yet it may be facilitated through statistical learning and acquired in a later phase (Polka et al., 2001). However, some questions about statistical learning remain. It is unknown how much statistical input is necessary for infants to build up phonetic categories in natural environment. Besides, the effect of duration (shorter or longer exposure) of the training phase in experimental conditions remains unclear. The statistical training phase in Maye's studies takes less than 3 minutes, showing how powerful and influential statistical learning can be. Moreover, the relationship between absolute frequency (ambient language input) and relative frequency (the proportion of each category) is not revealed from Maye's study. Finally, a pure distributional model cannot (at least fully) account for infants' generalization ability in the course of language acquisition. Some abstraction and rule-based learning abilities can also be tracked by infants before 12 months (Gómez & Gerken, 1999; Marcus, Vijayan, Rao & Vishton, 1999), though a certain amount of statistical information would be a prerequisite. Adding generalization to statistical learning may be a better way of explaining infant language acquisition rather than statistical learning per se (Adriaans & Kager, 2010).

1.4 Lexical tones in Mandarin Chinese

Different tones distinguish lexical or grammatical meanings in a tonal language (Wang, 1973). Four overt tones – flat, rising, dipping and falling – appear in Mandarin Chinese, marked in numbers from 1 to 4. Tone 1 (T1) is a high-level tone, while tone 4 (T4) is high-falling. In this language, each word is mono-syllabic and bears one tone. Same syllables with different tones form lexically contrastive minimal pairs. Some real words and sentences of the language are listed below.

Mandarin Chinese "ma"			
Chinese	Tone	Description	English
妈	ma1	high level	mother
麻	ma2	high rising	hemp
马	ma3	dipping	horse
骂	ma4	high falling	scold

Table 2 An illustration of lexically contrastive minimal pairs in Mandarin Chinese

Sentence 1:

Pinyin: Jìjì jíjí jī jī jì. (In tonal number mark: Ji4ji4 ji2ji2 ji1ji1 ji4.)

Chinese gloss: Ji-princess hurriedly beat chicken story

English: "the story of princess Ji beating a chick in a rush"

Sentence 2:

Pinyin: shīshì shíshí shí shī shǐ. (In tonal number mark: shi1shi4 shi2shi2 shi2shi1 shi3.)

Chinese gloss: Shi-female often eat lion history

English: "the history that Madam Shi often eats lions"

Tones can be seen as phonetic distinctions attached to the syllable at a prosodic level. Their properties include fundamental frequency (F0), amplitude (intensity) and duration[3] (Coster & Kratochvil, 1984; Halle, Chang & Best, 2004; Kong, 1987; Whalen & Xu, 1992). Tones and vowels create distinct vowel quality since the main organ used for production of tones (larynx) is different from vowels (oral cavity). An illustration of duration and fundamental frequency of tone in Mandarin Chinese can be seen in FIGURE 3.

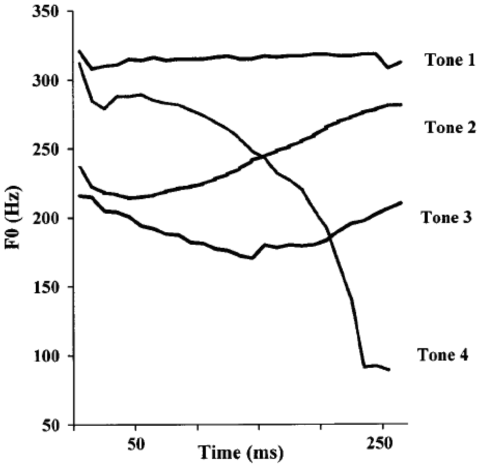


FIGURE 3 F0 contours of four tones in Mandarin Chinese

Source: Wang, Jongman & Sereno (2001)

Previous ERP, PET and fMRI studies suggest that lexical tones are perceived in different domains in the brain by native and non-native tone-language listeners. They seem to be perceived as linguistic information among native listeners and are processed in the left hemisphere, just as other speech segments (e.g. vowels) (Brown-Schmidt & Canseco-Gonzalez, 2004; Gandour, Dzemidzic, Wong, Lowe, Tong, Hsieh & Lurito, 2001; Hsieh, Gandour, Wong, & Hutchins, 2000; Kaan, Barkley, Bao

---

[3] Since the acquisition of prosody has been suggested to follow an order of fundamental frequency – timing – amplitude (Levitt, 1993), it is interesting to ask if the same pattern applies to tones. This is not discussed in this paper.

& Wayland, 2008; Klein, Zatorre, Milner, & Zhao, 2001; Wang et al., 2001), while by non-native listeners of a tonal language, they seem to be perceived as non-linguistic melodic information (Francis et al., 2008; Van Lancker & Fromkin, 1973; Wang, Behne, Jongman & Sereno, 2004). Native speakers of a non-tonal language show a degree of sensitivity to supra-segmental information such as stress and intonation (Repp & Lin, 1990; Lee & Nusbaum, 1993).

Interestingly, although standard Dutch does not have linguistic tones, Limburg dialect, spoken in the southern Netherlands, use tonal contrast to differentiate meaning (Mennen, Levelt & Gerrits, 2006). The two tones may be used to distinguish lexical or functional (e.g. singular or plural) meaning. Since the purpose of the current experiment is to test the tonal perception of infants from a non-tonal language background, infant subjects whose parents speak Limburg dialects have not been included in the study.

1.5 The acquisition of tones

More than half of the world's languages are tonal languages. Some early studies addressed the issue of early perception and production of tones in infants of a tonal language. Harrison (2000) tested 6–8-month-old Yorùbá and English infants on their perception of Yorùbá tones. The visually reinforced infant speech discrimination paradigm was used to test infants' sensitivity towards pitch changes. In the training stage, infants were trained to understand that sound changes triggered visual stimuli, while their head turns towards the visual stimuli given the shift or maintenance of speech sounds were measured in the testing stage. Results showed that Yorùbá infants were more sensitive to pitch changes on isolated syllables compared to English infants. Moreover, Yorùbá infants showed bias in hearing change at a single point in the testing frequency continuum. This suggested a language-specific categorical perception effect at that age only in infants from a tonal language.

Tse (1978) conducted a longitudinal study following one Cantonese speaking child's perception and production of tones from birth to 32 months of age. The perception of tone contrasts was judged by the child's association between spoken words and objects or actions when his father spoke some sentences to him in the real life. Results showed that perceptual discrimination of lexical tones started as early as 10 months of age. The high–level and low–level tones were the first to be acquired by the subject, followed by the high-rising and the mid-level tone. The low–rising tone was the most difficult to acquire in Cantonese.

With respect to the acquisition of tones in Mandarin Chinese, most studies focused on children of a later age range. Wong (2005) used picture-pointing and picture-naming tasks to test 3-year-old Chinese toddlers' tonal perception and production. Results showed that Mandarin toddlers perceived level (T1), rising (T2) and falling (T4) tones accurately whereas the dipping tone (T3) was perceived with difficulty. This result suggested a different developmental pattern for different tones, with an especially late command of the dipping tone. This claim was supported by previous studies (Clumeck, 1977; Li & Thompson, 1977). Wong also claimed that no correlation appeared between these toddlers' perception and production of tone.

Li & Thompson (1977) tested 17 Mandarin speaking children from 18 to 36 months on a picture-naming task. One major finding was that falling tones are produced earlier than rising tones. The authors attributed this finding to relative ease of learning to control glottal pitch and relative difficulty of rising pitch and claimed that falling tones require less physiological effort in production. However, Tuaycharoen (1977) studied one Thai infant's tonal development from 3 to 18 months of age and found that the mid-level and low-level tones were acquired in Thai at 12 months of age, followed by the rising tone. Hence, the rising tone emerged before the falling tone in Thai, showing variations of acquisition in different tonal languages. Clumeck (1977; 1980) traced tonal performance of one Mandarin child from 14 to 32 months of age and found that the child first discriminated rising and falling tones to differentiate needs and objects. Clumeck claimed that the order of perceptual competence with particular tonal contrasts was based mainly on the phonetic distinctiveness of the tone pair, and on the degree to which the tones are distinct despite variations resulting from tone sandhi[4] rules.

In a recent study, Zhu & Dodd (2000) tested Mandarin children's phonological acquisition (age

---

[4] Tone sandhi is the tonal change occurring when different tones come together in a word or phrase in some languages.

range: 1;6 to 4;6). Results showed that tones were acquired earlier than vowels and syllable-final consonants and that tone errors were rare yet the ‘weak stress’, used in rhotacism marking[5], was not commanded by the oldest child tested. Zhu & Dodd attributed the early discrimination and accurate usage of lexical tones to the fact that tones were used in distinguishing lexical meanings.

Previous literature regarding the acquisition of tones tend to use longitudinal study to track a small number of subjects for a whole picture of the tone acquisition, and focus more on production than perception. This narrowly affects the results of acquisition of Mandarin tones. Moreover, very few studies have looked into the beginning of tone perception at an early stage in an experimental way. This study will address the issue by using a tonal contrast in Mandarin Chinese.

1.6 Summary of introduction

In summary, infants use statistical learning to not only track frequency of sounds from ambient language input but also segment speech stream and acquire native phonemic categories. This ability may be one of the factors that result in a perceptual change which facilitates their acquisition of the native language. In this change, infants lose their sensitivity towards the speech properties that are irrelevant to the native phonemic inventory.

No previous study has addressed statistical learning of linguistic tones in infants. It is not only important to understand how infants learning a tonal language acquire native tonal categories, but also crucial to see whether non-tonal language infants are able to acquire or at least discriminate tones based on statistical information given in the training phase in the experiment.

No previous study has combined statistical learning with PR in infants. Maye’s study shows that after being trained with uni/bimodal frequency distributions, 6–8-month-old infants form respective perceptual sensitivity gained from the distributional input. However, the age ranges of both participant groups fall within the tonal PR time window (6 to 9 months). The study of interaction

---

[5] In Mandarin Chinese, rhotacism is the conversion of a consonant to a rhotic consonant (plus /r/) in a certain environment.

between statistical learning and PR is important because the potential influence of statistical learning in different phases of PR may reveal the plasticity of this perceptual change, i.e. whether PR can be compromised or even “revived” based on distributional input. It is interesting to find out how statistical learning and PR are weighted in an infant’s brain in the developmental course of the first two years of life. The current research may reveal the underlying mechanism of PR. Also it will look into whether PR puts an absolute limit in the developmental course and forms a linear decline in sensitivity to the non-native contrasts, or it is fairly flexible, with spaces for input probability and in the current case, allowing the interference of statistical learning.

Tone is a peculiarly interesting target to research because of its controversial PR time window. If tone perception follows the prosodic PR pattern, the onset of tonal PR should emerge at least earlier than 6 months when the vowel PR takes place, contradictory to the previous findings. It is thus necessary to look at tonal PR from different angles by using contrasts distinct from previous study.

The goal of the present thesis is multi-fold. First, it aims at answering whether infants are capable of using statistical learning on a non-native tonal contrast, that is to say, if statistical learning would facilitate or interfere with infants’ discrimination of a non-native tonal contrast. Second, and more importantly, this study intends to reveal the relationship between PR and statistical learning. Last but not least, the time window of tonal PR and its plasticity will be observed through the current study.

2 Research questions and hypotheses

The research questions of this study are:

1) Does statistical learning facilitate or disrupt infants' discrimination of a non-native tonal contrast?

2) Is this ability affected by tonal PR?

3) When does tonal PR take place?

The hypotheses are:

1) Infants are able to track the statistical information in the speech input. Infants familiarized with a bimodal distribution of tonal information will better discriminate the tonal contrast during the test phase than infants familiarized with a unimodal distribution. This is because a bimodal distribution helps infants build up two categories which eventually facilitate the contrast discrimination while a unimodal distribution weakens the contrast and eventually one category is built.

2) The statistical learning ability will be influenced by PR at least in the experiments, reflected in the progressive loss of sensitivity to a non-native contrasts. Younger infants before and in the beginning phase of PR are predicted to show statistical learning effect probably in the unimodal condition in which statistical input may decrease the discrimination ability whereas elder infants after or in the end phase of PR may either not show this effect or show the effect in the bimodal condition in which distributional information may increase the discrimination ability. One factor which may influence this possibility is the weighting between statistical learning and PR in the brain. A different result will show up, for example, if PR overrules statistical learning in the period of perceptual shift. Infants' discrimination patterns will be determined by their position in various PR phases.

3) The PR of tones occurs between 6 to 9 months according to previous studies by Mattock and his colleagues. The results from the current experiments will testify the validity of this time window.

3 Experiment 1

This experiment investigates whether infants' statistical sensitivity may occur at a prosodic level at a young age, as reflected in their discrimination of lexical tones (i.e. fundamental frequency (F0) value). If infants exposed to a bimodal distribution discriminate the non-native tone contrast while those trained in unimodal do not show preference, then it is likely that statistical learning applies to lexical tone acquisition.

3.1.1 Methods

3.1.2 Participants

Thirty-seven Dutch infants from 5 to 6 months (mean age: 169 days, ranging from 147 to 209 days) participated in the study (19 male). All infants were raised in monolingual Dutch families living in Utrecht, the Netherlands. The families reported that the infants' main language exposure in the first year of life was Dutch. The parents also reported that no hearing problem was found in their infants. Eventually data of 24 infants (mean age: 166 days, ranging from 147 to 181 days) were incorporated into a later analysis, with a drop-out rate of 35%. The exclusion criteria in this experiment were: age too old (2); crying (3) or fussing (2); failure to habituate after 25 trials in the test phase (3); too short LT (<2 s) on both change trials (1); dishabituation that differed by more than 2 standard deviations (SD) from the mean LT (2). Parental consent was obtained for each test. The infants were randomly divided into two groups, corresponding to two familiarization conditions (unimodal and bimodal), with 12 infants per group. Both groups were gender-balanced (5 males and 7 females in each group).

Maye et al. (2002) tested 8-month-old infants in consonant (VOT) contrast, prior to the PR timeline of consonant (10–12 months). In this study, it was reasonable to choose the age range prior to the PR of tones. Mattock & Burnham (2006) claimed that PR for tones occurs between 6 to 9 months. This time window was quite long and so far no other studies have provided other ranges. Considering that in the beginning phase of PR the discrimination ability to non-native contrast would at least be better than that in the end, it was decided to test infants of 5-6 months.

3.1.3 Stimuli

The stimuli for the unimodal and bimodal groups were four continua of the /ta1/–/ta4/ tonal contrast occurring in Mandarin Chinese (T1–T4). T3 was avoided due to the late acquisition and hence potential difficulty in discrimination among younger infants. The sounds were repeated and recorded by a Mandarin Chinese female speaker. Stimuli were created as follows: the natural T1 and natural T4 sounds were analysed in PRAAT (Boersma & Weenink, 2009). Four points on the pitch tier of both T1 and T4 were chosen: the starting point, the end point and two interpolate points (see dots in FIGURE 4). For each of two corresponding points, e.g. both starting points of T1 and T4, F0 was measured and the difference in F0-values was divided into 7 even portions (FIGURE 5). This was done for all four points on the pitch tiers of T1 and T4. The six intermediate values of all four points were used to create the intermediate stimuli between T1 and T4, thus creating a continuum of 8 stimuli (see Appendix II). This method was used to create a total of four continua. The other acoustic properties of the stimuli were kept to resemble normal speech sounds in Mandarin Chinese. Creating more than one continuum of the same tonal contrast kept the acoustic variation within the speaker, which is another way of resembling normal speech sounds. The average duration of the stimuli was 412 ms (ranging from 376 to 448 ms). Five native speakers of Mandarin Chinese judged all stimuli from the continua as sounding like normal tonal speech.

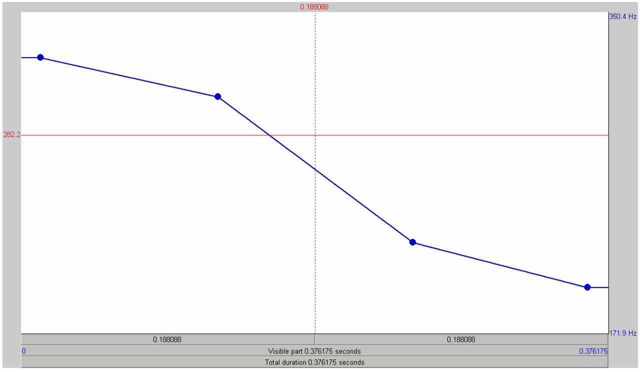


FIGURE 4 Visible pitch tier of stimulus 6 in Continuum A; dots representing the 4 points (Picture generated from PRAAT)

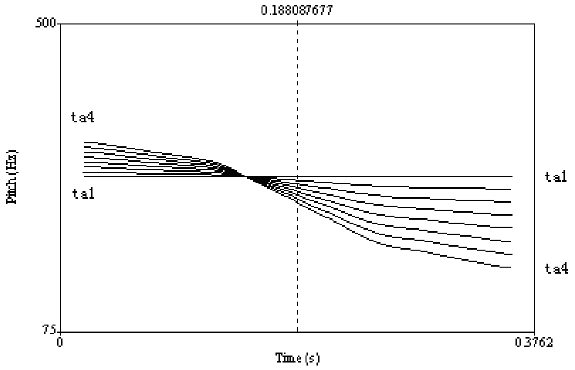


FIGURE 5 8 visible pitch contours along Continuum A

In FIGURE 5, /ta1/ is the flattest curve while /ta4/ is the sloppiest curve. The other curves correspond to the stimuli that are created along the continuum.

3.1.4 Apparatus

All infants were tested in the baby laboratory of Uil-OTS, Utrecht University. As can be seen in FIGURE 6, infants were seated on their parent's lap in a booth throughout the experiment. The booth was in a room that was sound-attenuated and a monitor was placed in front of the infants with a speaker hidden behind the textile of the booth. The monitor presented visual stimuli corresponding to the experimental design. The main function of the visual stimuli was to draw infants' attention at an appropriate time without affecting the validity of the observed data. No other appliance or any potential distraction to the infants was present in the booth. The parent wore a headphone throughout the experiment and masking music was played over the headphone so that the parent would not interfere with his or her infant's behavior. The tester could also talk to the parents over the headphone if necessary during the experiment.

FIGURE 6 Babylab testing environment

3.1.5 Procedure

3.1.5.1 General procedure

Following Capel (2008) and Maye et al. (2008), the experiment consisted of three phases. Infants were first familiarized with a unimodal or a bimodal frequency distribution in the familiarization phase; then in the test phase a habituation–dishabituation paradigm was used to test their discrimination ability towards certain stimuli. After that, infants' general unresponsiveness was checked by a post-test phase. In the end, the program ended with a happy song for the infants[6]. The Head-turn Preference Procedure (Nelson, Jusczyk, Mandel, Myers, Turk & Gerken, 1995) was used in all phases[7]: the looking time (LT) of the infants towards the monitor was recorded as the dependent variable that reflected the infants' attention to the auditory stimulus. The recording was conducted by a tester in the adjacent room, who observed whether or not the infants were looking at the monitor through a closed TV circuit. A button box was used by the tester to record the observation. This box was linked to a computer in which a correspondent tailor-made computer program (Veenker 2007) was running for the presentation of the auditory and visual stimuli. Both parents and the tester were blind about when the criterion from habituation to dishabituation was reached in the test phase.

---

[6] "Alle eendjes zwemmen in het water."

[7] The standard HPP has sounds playing from loud speakers on the sides of the booth as well. In this experiment the sound is solely from the front where the screen locates.

3.1.5.2 Familiarization phase

Infants were randomly assigned to one of the two conditions differing only in the relative frequency in the displayed stimuli in the familiarization phase. Twelve infants were exposed to the stimuli from the continuum presented based on a bimodal condition, while 12 infants were familiarized with the same stimuli with a different unimodal frequency distribution (FIGURE 7).

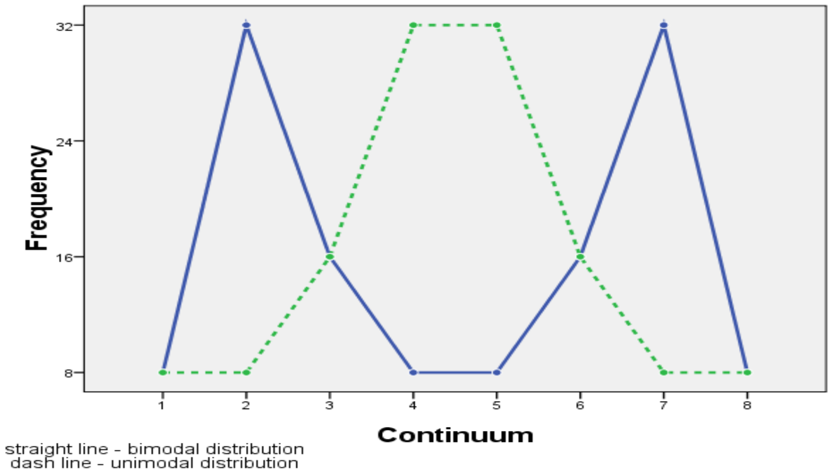


FIGURE 7 Unimodal and bimodal frequency distributions

The frequency of the eight stimuli for unimodal (from stimulus 1 to 8) was respectively 8-8-16-32-32-16-8-8, from which typical central peak (stimuli 4 and 5) could directly be observed. The frequency for correspondent bimodal was 8-32-16-8-8-16-32-8, from which two peaks appeared near the endpoints of the continuum (stimuli 2 and 7). In both conditions, the stimuli presented to the infants were equal in number (128 stimuli in total). These stimuli were presented in random order. The inter-stimulus interval was set as 1 s. The total duration of the familiarization phase was exactly 3 min.

Along with the auditory stimuli, a dynamic visual stimulus was presented on the monitor to keep infants' attention. The visual stimulus consisted of a continuously changing set of three colored pictures from a total set of 25 pictures within an invisible 3*3 grid on the screen (FIGURE 8). The goal of the visual stimulus was expected to lead to a lower drop-out rate in the familiarization phase. This visual stimulus was presented throughout the familiarization phase.

FIGURE 8 Instances of the dynamic visual stimulus

Infants of 5–6 months were generally attentive in the familiarization phase. No requirement was set as an exclusion criterion with respect to the LT in this phase. However, infants who were fussy or crying in this phase were excluded from the study, as in other phases.

3.1.5.3 Test phase

Right after the familiarization phase, the test phase began. Infants were tested on their discrimination of stimuli 6 and 3. The test phase consisted of a habituation and a dishabituation stage. In the habituation stage, stimulus 6 was presented to the infants repeatedly with an inter-stimulus interval of 1 s. This repeated presentation was set as a trial. A trial stopped when a criterion for each of two conditions was met: the infants looked away for more than 2 s, or the stimuli repetition continued for a maximum of 45 times. In the first case, the trial was considered ineffective and was dropped from total effective trials. A new trial immediately started after the former trial. The starting of a trial was controlled by the tester to make sure that infants were attentive in the beginning.

Three consecutive effective trials formed a “window”. The first three effective trials in the habituation stage were called window A. The number of trials in the habituation phase depended on a habituation criterion: 65% of the total LT of window A. Starting from the fourth trial, if the LT of any three consecutive effective trials (named as window B) decreased to at least this criterion, the habituation stage ended and the dishabituation stage started. FIGURE 9 illustrates an example. Suppose the total LT for the first three effective trials were 10 s, formed window A and set the habituation criterion as 65%*10 s = 6.5 s. In the first (trial 4–6) and second (trial 5–7) window B candidates, the total LT was 8 s and 7 s respectively, larger than the criterion. The total LT for the third candidate of window B was 6 s, smaller than the 6.5 s criterion. The criterion was met.

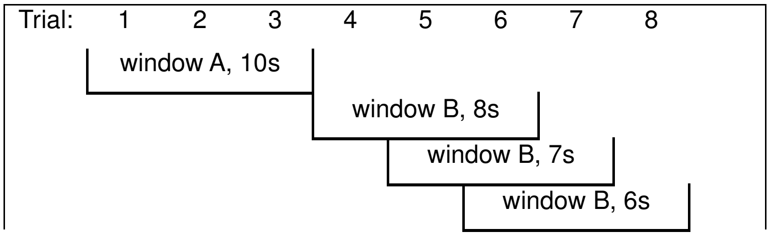


FIGURE 9 An example showing the the habituation criterion paradigm

In the most rapid case, at least 6 trials were required for the habituation stage if they were all effective (no less than 2 s). Maximally 25 trials were set as the upper limit of the habituation stage. An increase in habituation criterion (from 50% in previous studies to 65% in this study) was expected to cause a decreased drop-out rate, since 35% shorter attention is faster to reach than the original 50% criterion.

The number of trials in the dishabituation stage was always set as 2, regardless of the effectiveness of the LT in each trial. In both trials, stimulus 3 was presented repeatedly. A trial ended when the infants looked away for more than 2 s, or the stimuli repetition continued for a maximum of 45 times. Data of infants who looked less than 2 s to both of the dishabituation trials were excluded from the analysis. Note that stimulus 6 and 3 were selected in the test phase due to their equal distributional frequency in both unimodal and bimodal in the familiarization phase despite the different frequency of other stimuli between the frequency distributions. This excluded the possibility that attributed the difference in LT to the unequal frequency during familiarization between the two conditions.

Along with the auditory stimuli, a static visual stimulus, colorful bull's eye (FIGURE 10) was presented on the monitor. Compared to the dynamic visual stimulus in the familiarization phase, this less interesting stimulus was expected to reduce infants' awareness on the visual stimulus and to ensure that infants' LT relied on only the auditory stimuli in the test phase.

FIGURE 10 A static visual stimulus - colorful bull's eye

Infants' discrimination abilities were revealed by comparing the mean LT of the last two effective habituation trials and the two effective dishabituation trials. Bimodal group infants were expected to show more discrimination of stimuli through significantly longer LT in the dishabituation trials as a novelty effect while unimodal group infants should show no such preference. Table 3 illustrated another example of test phase.

	Habituation						Dishabituation	
Trial:	1	2	3	4	5	6	1	2
	20s	10s	6s	4s	3s	3s	10s	6s

Table 3 The habituation–dishabituation trials showing infants' discrimination ability

3.1.5.4 Post-test phase

The post-test phase immediately started when the test phase ended. It was set as a test of general unresponsiveness of the infants. This phase consisted of one repeated stimulus only - a Chinese sentence "ni3 zhen1 xing2 a4" (Translation: "You do really well!"). This sentence contained four mono-syllabic words, each with a distinct tone. The sentence was spoken by a female voice originally from the same speaker of the previous stimuli. The visual stimulus from the test phase (the colorful bull's eye) remained unchanged in post-test phase. This phase ended when the infants looked away for more than 2 s. Infants who did not look longer to the post-test trial than to the dishabituation trials, would be excluded from analysis for the reason of unresponsiveness, a criterion which was never met throughout the experiment[8]. In the end, an infant-friendly song was

---

[8] Normally infants showed great interest to the new complex stimulus in the post-test after a relatively boring

played, providing infants with a nice mood when they left the booth. The whole experiment ended.

3.2 Results

3.2.1 Habituation

The mean LT for windows A and B, representing the first three and last three effective trials respectively in the habituation phase, are compared by repeated measures ANOVA. The between-subjects factor is two-level familiarization condition (unimodal and bimodal) while the within-subjects factor is the LT for windows A and B. The main result is shown in FIGURE 11.

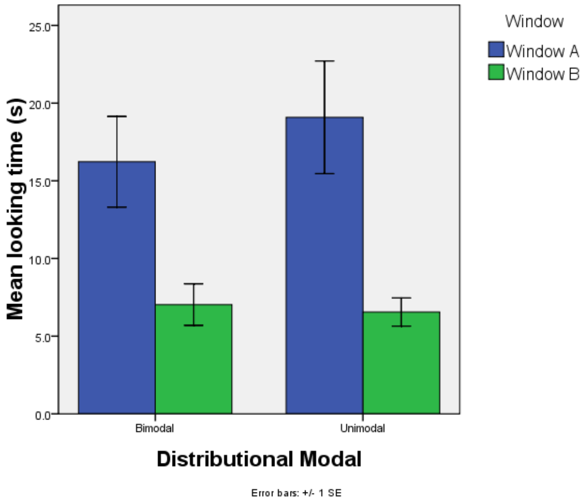


FIGURE 11 LT shift between window A and window B

A significant difference is obtained between the two windows, F (1, 22) = 37.008, p < 0.001. This is not unexpected since the LT reaching point of window B is 65% duration of window A. The key point here is that no interaction between window and modal (uni/bi) is observed, F (1, 22) = 0.874, p = 0.360. This indicates that there is no significant difference in LT between the two frequency distributions when the window shifts from A to B, hence infants under both conditions are habituated.

---

habituation and dishabituation phase. There were however two infants who were shocked by the sudden change of the stimulus and began to cry. Immediately, the tester ended the post-test phase and started the infant-friendly song to keep infants happy. Infants' data were not excluded under such circumstances.

Other potential factors are compared through Univariate ANOVA. These factors include the number of effective trials and total LT in habituation, etc. These factors might be relevant because they reflect the actual duration in LT of stimulus 6 in habituation phase. Results reveal no significant difference across these factors between the two modal groups (Number of effective trials: F (1, 22) = 0.116, p = 0.737; total LT in habituation: F (1, 22) = 0.033, p = 0.857).

3.2.2 Habituation and dishabituation

The mean LT for last 2 habituation trials and 2 dishabituation trials are compared by repeated measures ANOVA. The between-subjects factor is the familiarization condition, set as two levels (uni- and bimodal) while the within-subjects factor is mean values of 2 last habituation and 2 dishabituation trials.

Recalling from section 2.1.4.3, stimuli 6 and 3 are used in the habituation and dishabituation phase respectively because they are played in equal frequency in either distributional modal. Since all other settings are mirrored in the two groups, infants' potential response difference should be interpreted as caused by the different exposure of frequency distributions during familiarization.

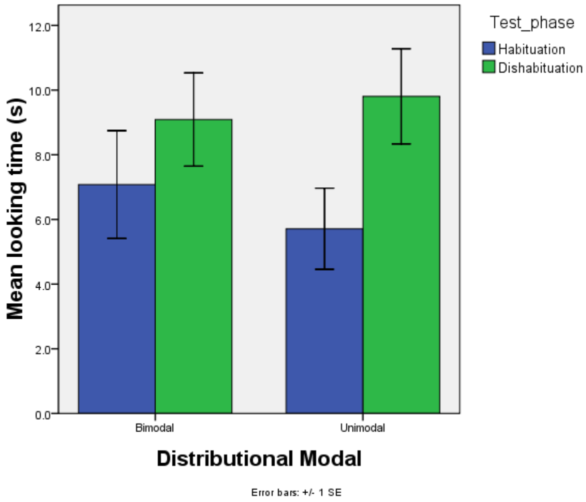


FIGURE 12 LT shift from habituation to dishabituation phase

As FIGURE 12 shows, a significant difference is obtained between LT in habituation and dishabituation phase, F (1, 22) = 6.254, p = 0.02, indicating infants' sensitivity towards the tonal change. No interaction between phase change and modal (uni/bi) is observed, F (1, 22) = 0.728, p = 0.403. This means that there is no significant difference in LT between the two frequency distributions from habituation trials to change trials. In summary, infants from both modal groups catch the difference.

3.3 Discussion

The finding for the habituation phase is consistent with previous studies. By all means, the criterion is set as a 65% decrease from window A to B. The similar performance between the two training groups reveals that the frequency distributions do not strongly affect the habituation procedure — in the end subjects all become bored with the stimuli in the same fashion, independent of the number of effective trials or the total LT in habituation phase. The key finding lies in that infants from both frequency distributions discriminate the non-native tonal contrasts. This seems contradictory to the previous prediction of statistical learning. In the study by Maye and colleagues (2002) and Capel (2008), bimodal group infants' sensitivity is significantly stronger in the bimodal than in the unimodal condition. What can account for the current finding?

The first and crucial factor lies in the different age ranges of the subjects, which differ greatly across studies (Table 4).

Studies		Infant age		Type of Contrast
		in months	in days	
Current study	2010	5–6	147-181	tone
Maye et al.	2002	6–8	181-288	consonant
Capel	2008	10–11	290-348	consonant

Table 4 Different age group and type of contrasts in statistical learning studies

If statistical learning (relative frequency) is accessible for infants of 5–6 months, the results of this study show a ceiling effect. It is very likely that the initial tonal sensitivity, possibly an innate boundary specification, overrules statistical learning and inhibits its appearance. The contrast could be too obvious for 5-month-old infants for statistical learning to even be relevant. Note that in Maye's study infants of 6–8 months are tested. Since previous studies suggest that tonal PR occurs from 6 to 9 months (Mattock & Burnham, 2006; Mattock et al., 2008), the tonal PR onset is reached in Maye's study but not reached in the current study. Hence, the finding that infants are sensitive to tones in 5 months regardless of the modal is likely to be a ceiling effect of general perception. General speech sensitivity to tones in the early ages of infants overrules statistical learning at least in the experimental settings.

Another possibility, though quite unlikely, is that the current type of statistical learning requires a certain degree of maturation in the brain. In fact, no statistical learning in the language domain has been shown for infants before 6 months, the age at which infants are still open to contrasts across languages. The memory load or cognitive demands for 3 minutes of continuous exposure to statistical information is perhaps not what infants can achieve at 5 months of age. Alternatively, it could be that infants before 6 months are not capable of using certain type of statistical learning mechanism, such as recognizing the absolute frequency of the tonal input while ignoring the relative frequency in the current experiment.

Comparing the previous studies, one last possible factor would be the influence of the type of contrast: suprasegmental features are not applied to statistical learning while segmental features do. This logically possible factor does not stand, as can been seen in the next experiment. By all means, in this experiment statistical learning might be affected by infants' tonal PR.

Tracing back to the stimuli design, within-speaker variations were added into the input stimuli in order to create a more natural and linguistic perception to the infants (The speaker pronounces /ta1/–/ta4/ four times.). This may lead to another possible situation: infants cannot integrate the variations into one gross tonal map, possibly due to the immature brain capacity or the given high cognitive demands, and hence not able to represent the whole statistical category. This will predict that reducing the variability (i.e. by using one sound) may lead to a change in the results for the unimodal group. Further studies should address this issue.

Variations are observed among subjects as well. The variation can be seen as reflecting individual degree of sensitivity towards tones. Within 12 subjects in each modal group, 7 infants show a strong novelty effect while other 5 do not seem to be sensitive to tones. Although infants' experiment have larger variation than adults' experiment on the perception of tones, it is important to see that more than half of infants actually look longer to the shifted stimuli in each modal, indicating their high sensitivities are not likely to be caused by chance performance. It is interesting to see if infants who are initially not sensitive to the tonal contrast will become sensitive at some point. Possibly, some infants are generally slower in sound perception than others at early ages. A longitudinal study with an adequate number of subjects may provide an answer to this question.

4 Experiment 2

The previous experiment provides a picture of 5–6-month-olds Dutch infants' tonal perception under a distributional model. The research questions are not yet fully covered: no sign shows a facilitation effect for these infants' discrimination of a non-native tonal contrast through statistical learning. Moreover, although Experiment 1 looks into the onset of tonal PR, the offset of tonal PR has not been verified. According to Mattock et al. (2008), 11–12-month-old infants have passed tonal PR stage. This may predict that these infants will lose their sensitivity to tones and show no trace of statistical learning. Experiment 2 studies 11–12-months-old Dutch infants' perception under the same condition.

4.1 Methods

Thirty-eight Dutch infants from 11 to 12 months (mean age: 357 days, ranging from 339 to 393 days) participated in the study (22 male). The same criteria were used in participant selection with respect to language exposure and health/hearing condition. Eventually data of 28 infants (mean age: 356 days, ranging from 334 to 387 days) were included into later analysis, with a drop-out rate of 26%. The exclusion criteria in this experiment were: age too old (2); failure to habituate after 25 trials in the test phase (1); too short LT (<2 s) on both change trials (1); fussing or LT not reaching 60% of the total LT in the familiarization phase (6). The infants were randomly divided into two groups, corresponding to two familiarization conditions (unimodal and bimodal), with 14 infants per group. The same stimuli, apparatus and whole testing procedure were used in the experiment.

Setting LT not reaching 60% of the total LT as one of the criteria not only provided one way of measuring fussiness, but also ensured that infants were paying enough attention to the familiarized stimuli in order to be "soaked" into the distributional information hidden within. In Experiment 1, 5–6-month-old infants were generally more attentive and no such requirement was set.

4.2 Results

4.2.1 Habituation

The mean LT for windows A and B, representing the first three and last three effective trials respectively in the habituation phase, are compared by repeated measures ANOVA. The between-subjects factor is two-level familiarization condition (unimodal and bimodal) while the within-subjects factor is the LT for windows A and B. The main result is shown in FIGURE 13.

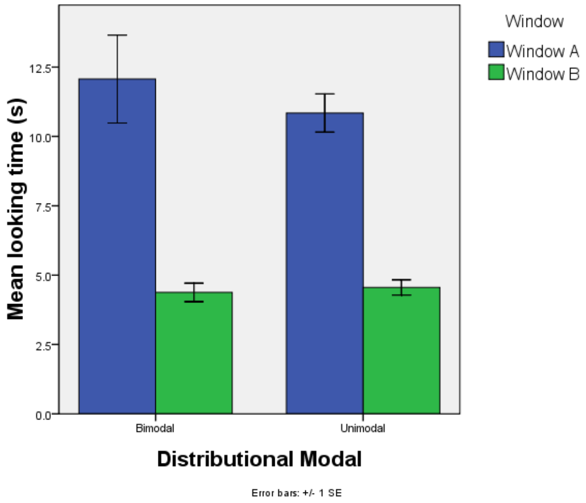


FIGURE 13 LT shift between window A and window B

A significant difference is obtained between the two windows, F (1, 26) = 68.558, p < .001. Again, no interaction between window and modal condition is observed, F (1, 26) = 0.690, p = .414. This indicates that there is no significant difference in LT between the two frequency distributions when the window shifts from A to B. Infants under both conditions are habituated.

Other potential factors are compared through Univariate ANOVA. These factors include LT in familiarization phase, the number of effective trials and total LT in habituation. Results reveal no significant difference across these factors between the two modal groups (total LT in familiarization: F (1, 26) = 0.768, p = .389; effective trials: F (1, 26) = 0.615, p = .440; total LT in habituation: F (1, 26) = 2.053, p = .164).

4.2.2 Habituation and dishabituation

The mean LT for the last 2 habituation trials and 2 dishabituation trials between the two modal groups are compared by repeated measures ANOVA. The between-subjects factor is two-level familiarization condition (unimodal and bimodal) while the within-subjects factor is the mean LT of 2 last habituation and 2 dishabituation trials.

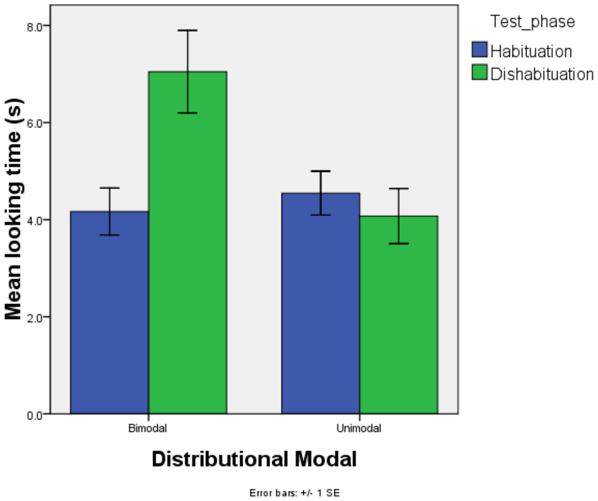
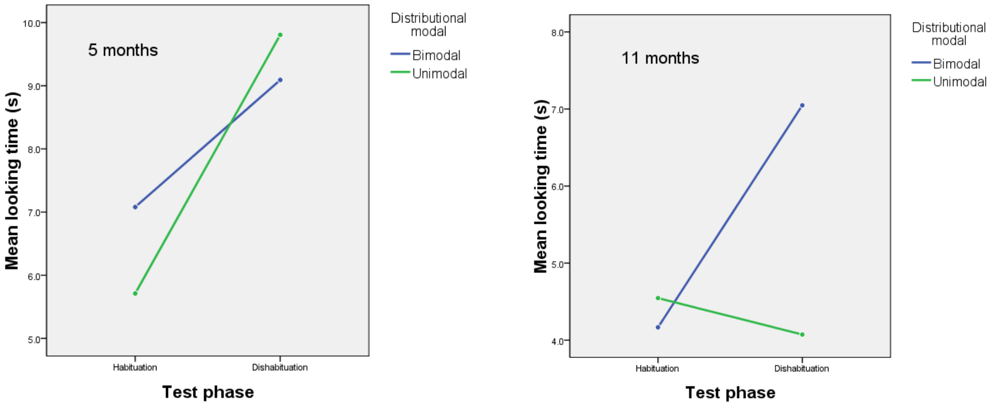


FIGURE 14 LT shift from habituation to dishabituation phase

A significant difference is obtained between LT in habituation and dishabituation phase, F (1, 26) = 4.542, p = .043. From FIGURE 14, it can be observed that this significance is mainly caused by the longer LT of infants under the bimodal in dishabituation phase. The key finding is that a significant interaction is observed between phase change and modal (uni/bimodal), F (1, 26) = 8.807, p = .006. This means in the dishabituation phase infants from the bimodal group look significantly longer than infants from the unimodal group when hearing the difference. A further paired sample t-test on unimodal group between the LT in habituation and dishabituation phase reveals no significant difference (t = 1.013, p = .330), indicating that infants in the unimodal condition are not sensitive to the tonal shift, contrary to their bimodal peers. In summary, only infants from the bimodal group catch the difference.

4.2.3 Cross-age comparison between infants of 5 and 11 months

By putting the results of both 5-month-old and 11-month-old infants' data into a repeated measures ANOVA model with an additional factor of age, more interesting results are revealed. Neither the age factor nor the modal factor yields significant difference (Age: F (1, 48) = 2.082, p = .156; modal: F (1, 48) = 0.245, p = .623). However, the interaction between these two factors is significant, F (1, 48) = 4.495, p = .039). Based on the previous analysis and FIGURES 15 & 16, it can be deduced that infants in the 11 months unimodal group behave significantly different from infants in other conditions.



FIGURES 15 & 16 LT shift in uni/bimodal condition across ages

4.3 Discussion

4.3.1 Eleven-month-old infants in the uni/bimodal condition

The finding in the habituation phase is as expected: both groups of infants are habituated in the end of this phase. The number of effective trials or total LT in familiarization or habituation does not yield a significant difference. The crucial asymmetric perception between the two modal groups in the tonal stimuli shift is consistent with previous studies (Capel, 2008; Maye et al., 2002, 2008). More interestingly, this experiment reveals the relationship between PR and statistical learning.

The testing subjects are 11-month-old infants. Two hypotheses emerge regarding the offset of tonal PR. Either way, the result shows a clear indication that statistical learning is interfering with tonal PR.

The first hypothesis is that the offset of tonal PR is later than 11 months, different from what has been suggested in previous studies of Mattock and colleagues. In this case, infants would possibly still keep some degree of sensitivity towards tones and either both modal groups or only bimodal group may notice the difference. In the current experiment, the fact that the unimodal group does not notice this distinction has to be explained by the statistical interference in the training phase, namely the unimodal exposure either does not change or it decreases infants' sensitivity to the tonal stimuli.

A further look into the first hypothesis reveals that tonal PR has not passed at 11 months and may be longer than what previous studies suggest. Given the high frequency of prosodic input in the ambient language such as pitch variation in intonation, it seems that tonal PR is supposed to occur at an early stage. The first explanation could be that tones are intrinsically distinct from other prosodic features. The second answer may be related to the relative psycho-acoustic saliency of the tonal contrast per se. Psycho-acoustically less salient tonal pairs may be perceived later than more salient ones. This goes with the tonal production data from native Chinese toddlers who acquire the most difficult tone (dipping, T3) the latest. However, in the current case it is difficult to evaluate how "psycho-acoustically salient" the tonal pair is so as to compare it to previous studies that use different tonal pairs (and probably in different languages). No unified method or standard has been proposed in the literature regarding the psycho-acoustic saliency of tones, and future studies should address this issue.

The second hypothesis is that the offset of tonal PR occurs earlier than 11 months, probably at 9 months as previous studies suggest. In that case, infants should no longer hold sensitivity to tones. The fact that bimodal group infants regain the sensitivity to tones has to be explained by the facilitation of the bimodal exposure in the familiarization phase, and the high plasticity in the offset of tonal PR.

Looking into the second hypothesis, it is likely that the onset and offset of tonal PR is fairly flexible. That means even though tonal PR occurs from 6 to 9 months of age, it could be extended to a longer range. As suggested by Werker & Tees (2005), PR can be seen as an "optimal period" with a plastic onset and offset time window, rather than a strict "critical period". If this claim is true, the current experiment can be seen as showing the plasticity of tonal PR triggered by bimodal distributional input. Similar finding supporting this flexibility of PR comes in the bilingual study by Bosch & Sebastián-Gallés (2003b). Catalan-Spanish simultaneous bilingual infants form a U-shape in their speech perception development in the first year when discriminating one Catalan vowel contrast. They distinguish the contrast at 4 months, fail to differentiate it at 8 months and at 12 months their discrimination ability comes back. This is explained by the frequency of input: in a simultaneous bilingual situation, the input of each language is reduced in half compared to monolingual peers. PR is thus delayed, as Bosch & Sebastián-Gallés argue, in bilingual infants. However, their data may also suggest that bilingual PR is not delayed but flexible in its onset and offset time window, following the "optimal period" hypothesis. Although Bosch & Sebastián-Gallés test a native vowel contrast while this study focuses on a non-native tonal one, the principle and mechanism behind PR is likely to be the same regarding its plasticity.

To distinguish the two hypotheses, a baseline study is needed to determine whether or not infants of 11 months of age can distinguish the same tonal contrast without the exposure of statistical interference. The result will indicate the PR situation of the tonal contrast at 11 months. Further studies should focus on this issue.

Similar to the first experiment, variations among infants are observed in this test and these 11-month-old infants are not totally "deaf" towards the tonal stimuli in the unimodal condition. Two infants indeed notice the shift of the sound.

4.3.2 Cross-age comparison between infants of 5 and 11 months

Regarding the 2*2 (age*modal) cross-age comparison between 5 and 11-month-old infants under different modal conditions, results show that only 11-month-old infants in the unimodal condition fail to discriminate the tonal contrast. This suggests that the result of statistical learning is not independent across age. On the contrary, it greatly depends on the PR of the subject and probably also the intrinsic saliency of the contrast tested. This potential influence of PR will be further tested in the next experiment.

The results between 5-month-old and 11-month-old infants raise an interesting issue, since no trace of statistical learning can been seen in the first experiment. It could be that in order for statistical learning to occur, at least a small amount of PR should have occurred. A paradox arises if native tonal contrast acquisition solely depends on statistical learning. If Chinese infants use the same learning process to acquire the contrast, their ability for statistical learning should also initially be overruled by their general speech contrast sensitivity as experiment 1 shows. However, it is highly unlikely that they pass through a stage of decreased tonal sensitivity to activate statistical learning. Moreover, if PR is interpreted as a result of (unimodal-like) statistical learning, then the puzzle can be solved. Statistical learning may function to (a) decrease sensitivity to non-native contrasts through a unimodal-like distribution if a native phoneme is either one of the members of the non-native contrasts or relatively familiar, (b) maintain and facilitate the sensitivity to native contrasts obtained from initial sensitivity through a bimodal frequency distribution, and (c) possibly reopen the gate of non-native contrasts in infants of 11 months who lost at least part of their sensitivity.

5 Experiment 3

Infants up to 11 months still distinguished the tonal contrast in the bimodal condition. To determine whether they were capable of doing so regardless of the age factor, infants of 14 months were tested to further understand the influential time range of statistical learning on tonal PR.

5.1 Methods

20 Dutch infants from 13 to 14 months (mean age: 437 days, ranging from 395 to 463 days) participated in the study (16 male). The same criteria were used in participant selection regarding language exposure and health/hearing condition. Eventually data of 14 infants (mean age: 436 days, ranging from 395 to 463 days) were included into later analysis, with a drop-out rate of 30%. The exclusion situation in this experiment were: age too young (1); crying (2), fussing (1) or LT not reaching 60% of the total LT in the familiarization phase (1); dishabituation that differed by more than 2 standard deviations (SD) from the mean LT (1). All infants participated in the bimodal condition, since the previous experiment had shown that infants in the unimodal condition no longer discriminate the tonal contrast and it was unlikely for these infants to recover from PR given the unimodal distribution. The same stimuli, apparatus and whole testing procedure were used in the experiment.

5.2 Results

5.2.1 Habituation

The mean LT for windows A and B, representing the first three and last three effective trials respectively in the habituation phase, are compared through a paired sample t-test. As can be seen in FIGURE 17, a significant difference is obtained between the two windows ($t = 5.057$, $df = 13$, $p < .001$), as an indication of habituation.

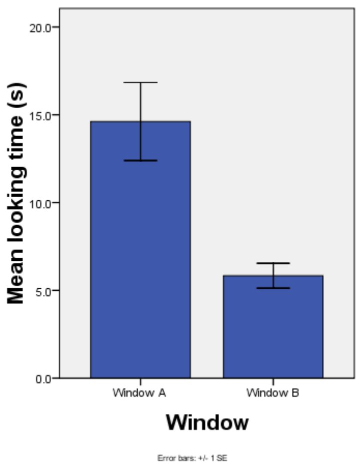


FIGURE 17 LT shift between window A and window B in bimodal condition

5.2.2 Habituation and dishabituation

The mean LT for the last 2 habituation trials and 2 dishabituation trials are compared by paired sample t-test. No significant difference is obtained (t = 1.351, df = 13, p = .200), showing that infants of 14 months are no longer attentive to the tonal shift in the stimuli (FIGURE 18).

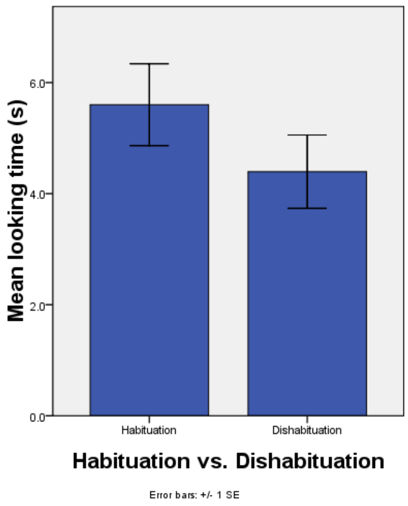


FIGURE 18 LT shift from habituation to dishabituation phase in bimodal condition

5.2.3 Cross-age comparison between infants of 11 and 14 months

Putting the results of both 11-month-old and 14-month-old infants' data in the bimodal condition into repeated measures ANOVA model with age as a between-subject factor reveals more interesting results. In terms of window shift, the results of two groups do not reach a significant level (F (1, 26) = 0.413, p = .526), indicating that there is no significant difference between the two age groups in habituation phase. Hence both groups of infants are habituated as expected.

The key finding is that a significant interaction is obtained between the LT shift from habituation to dishabituation phase and the age groups (F (1, 26) = 8.728, p = .007), revealing that 14-month-old infants under bimodal condition no longer discriminate the contrast (FIGURE 19).

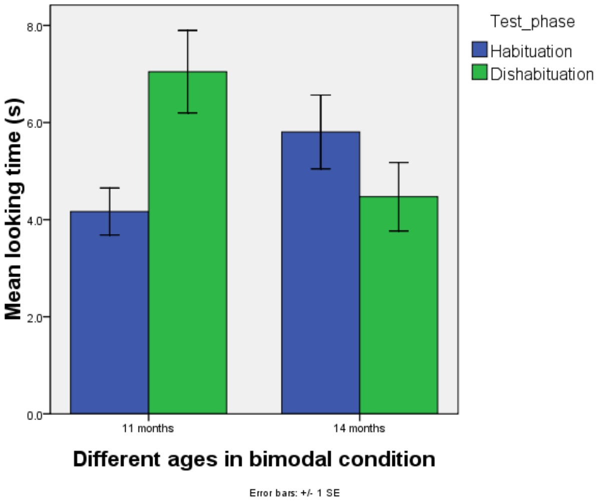


FIGURE 19 LT shift from habituation to dishabituation phase in bimodal condition between ages

Interestingly, a Repeated Measures ANOVA between 11-month unimodal group and 14-month bimodal group reveals a very similar pattern, with no significant difference in either LT of window shift in the habituation phase (F (1, 26) = 2.277, p = .143) or interaction between the age factor and the test phase (F (1, 26) = 0.623, p = .437). In other words, 14-month-old infants under bimodal condition seem to perform the same way as 11-month-old infants under unimodal condition.

5.3 Discussion

5.3.1 Fourteen-month-old infants in the bimodal condition

The results of Experiment 3 show that 14-month-old infants under bimodal condition do not distinguish the difference. The influence of statistical learning seems to be restricted to an earlier age range before 14 months, even in a bimodal condition. The first interpretation would be that the offset of tonal PR ends before 14 months and PR rules over statistical learning at the age of 14 months. Infants lose sensitivity to tones regardless of the provided tonal distribution reflected in their unawareness of the tonal shift.

Alternatively, it could be that the given type of statistical learning does not apply to infants of older ages. Similar to the claim that this might not apply to infants at early ages, this unlikely hypothesis needs to be verified by implementing the same distributional modal to children of an older age or even adults. The potential confound would be whether non-linguistic tonal perception might play a role even though older children or adults are able to perceive the distinction. In the current experiment, it is reasonable to exclude this possibility since no differences in LT are observed between habituation and dishabituation trials.

Finally, it has to be noted that infants' drop out rate in experiment 3 is higher than the other two experiments. A few 14-month-old infants have learned how to walk and are not willing to sit and go through the whole familiarization and habituation procedure. It could be that these infants do not notice the tonal shift simply because they are not paying attention. However, in the post-test phase when the new stimulus (a 4 syllable sentence) is played, these infants immediately pay attention to this change. Hence this possibility is unlikely to account for the finding of the experiment. Similar to the previous experiments, variations in behavior can be observed among infants of 14 months. For example, 1 infant indeed catches the difference when the phase changes.

5.3.2 Cross-age comparison among infants of all age-groups

In summary, three age groups and two frequency distributions are tested across groups. In total 5 groups appeared in the three experiments (Table 5). Pooling all the data into a repeated Measures ANOVA, a general picture is revealed.

	5 months	11 months	14 months
Bimodal	12	14	14
Unimodal	12	14	N/A

Table 5 Number of participants and groups across 3 experiments

In terms of window shift in habituation phase, no significant difference is observed regarding age groups (F (2, 61) = 2.194, p = .120), distributional modals (F (1, 61) = 0.263, p = .610) or the interaction between the two factors (F (1, 61) = 1.578, p = .214). This indicates that the habituation procedure works consistently across all age ranges and frequency distributions.

In the habituation-dishabituation phase, the results show that the age group is a significant factor (F (2, 61) = 4.315, p = .018), as well as the interaction between the age group and the frequency distributions (F (1, 61) = 4.864, p = .031). Meanwhile the distributional modal is not significant across groups (F (1, 61) = 0.265, p = .608).

Univariate ANOVA is used to set the 11 month bimodal group as a baseline in order to look at LT difference from habituation to dishabituation. Pairwise comparison shows that the 11 months bimodal group is not significantly different from the 5 months bimodal (p = .620) and the 5 months unimodal group (p = .489). However, it is significantly different from the 11 months unimodal group (p = .050) and the 14 months bimodal group (p = .014), while the latter two groups do not differ much (p = .608). This shows that the whole data can be briefly categorized into two parts, based on whether or not infants can catch the tonal shift in the experiment.

6 General discussion and conclusion

Three age groups are tested in the experiments. The results reveal a gross interaction between PR and statistical learning (Table 6): different statistical learning patterns are reflected in different stages of tonal PR. In short, it seems that statistical learning does not apply before the onset and after the late offset of PR. However, it does apply in the early offset, or probably during the PR time window.

Age	5 months	11 months	14 months
SL	No	Yes	No

Table 6 Brief summary of statistical learning effect across ages

Thus a general picture can be revealed: before the tonal PR, 5-month-old infants can well detect the tonal shift regardless of the statistical interference. During or shortly after tonal PR, 11-month-old infants are greatly influenced by the familiarized distributional information. Long after tonal PR, 14-month-old infants probably consider tones non-informative[9] and no longer pay attention.

At this point, the essence of PR and statistical learning should be reconsidered. PR and statistical learning could be a) two different mechanisms which interact with each other, or they are b) the same mechanism fundamentally.

a) According to the separate mechanism hypothesis, infants' perceptual sensitivity to language contrasts may naturally reduce and diminish, while statistical learning slows down or stops this process by receiving the bimodal distributional input from the ambient language and mapping this onto two perceptual categories. When such distributional input is lacking in the environment, the natural process continues and eventually those contrasts that are not saved by statistical learning will be lost. Based on this hypothesis, Table 6 can be interpreted as follows: at 5 months, tonal PR has not come into play and the initial sensitivity does not need support from statistical learning

---

[9] To most infants, the distinctions that are considered not useful in the native language environment will be filtered out (Werker, 1994).

mechanism. At 11 months, tonal PR is either in the process or has come to an end, and at this time statistical learning has the power to overrule it. At 14 months, tonal PR has completed and possibly becomes irreversible. In this interpretation, statistical learning has an influence on perceptual sensitivity, but only during a limited time window of PR, in which contrasts have already begun to decrease but not completely. Still, it is necessary to ask questions such as why PR is language-specific rather than domain-general if it is an independent mechanism; and what motivates two separate systems instead of one.

b) According to the unified mechanism hypothesis, PR is seen as a consequence of statistical learning. Exposure to a unimodal distribution (or no exposure, hereinafter) weakens sensitivity to the contrasts, while a bimodal exposure will reinforce and strengthen the initial sensitivity. This hypothesis is in line with Maye and her colleagues who suggest that statistical learning may be the underlying mechanism of perceptual changes in the second half of the first year of life. Initial sensitivity will be reduced due to a unimodal distributional exposure. It is likely that statistical learning also requires a certain level of cognitive maturation in order to be activated at around 6 months of age and as a result perceptual changes occurs only from that age onwards. Based on this hypothesis, Table 6 can be interpreted as follows: at 5 months, statistical learning has not become available and hence PR has not yet started. Infants show initial (universal) sensitivity to the tonal contrast. At 11 months, non-native tonal language infants can be considered as having no tonal exposure in the natural environment. However, since statistical learning is at work, infants are highly sensitive to distributions in the lab and their decreasing sensitivity can be “reversed” in the bimodal condition. At 14 months, due to the prolonged (lacking of) exposure to lexical tones in the ambient language, infants’ tonal sensitivity has passed beyond the point (PR) at which it can be reversible in the lab.

Moreover, the separated and unified mechanism hypotheses make different predictions on the reason why PR starts at around 6 months of age. In the former case, it could be due to a natural decline of perceptual abilities around 6 months, while in the latter case, it is because of the cognitive domain required for statistical learning matures at around 6 months of age. Intuitively, the unified hypothesis is likely to be favoured compared to the separated one. Further studies should work on this issue.

To answer the research questions: statistical learning does facilitate infants' discrimination of a non-native tonal contrast, but only within or near the offset of its PR time window. The time window of tonal PR is still unclear. It may take place from 6 to 11 months, longer than previously thought, and possibly due to the distinct perceptual saliency and difficulty of each contrast, or indeed from 6 to 9 months yet having an offset window with high plasticity.

As for the infants older than 14 months, it is suggested that their linguistic perception to tonal contrast further decay without a constant bimodal distribution in the ambient language. If there is discrimination to the contrast, it should not be considered as a linguistic effect. Rather, infants' statistical learning ability remains untouched cross-domain. Adult experiments (Appendix I) reveal that although Dutch grown-ups do not perceive lexical tone in a linguistic categorical fashion, they are not "deaf" to the tonal information and seem to be sensitive to the (non-linguistic) natural saliency of the tonal contrasts. Meanwhile, Chinese adults perceive tones more categorically than Dutch peers. The focus is on how lexical tones are perceived exactly in Dutch infants. The possibilities of a) non-linguistic perception and b) intonation perception both exist. Saffran and colleagues (Saffran, Johnson, Aslin & Newport, 1999) show that non-linguistic tone sequences can be segmented just as words by infants of 8 months based on statistical information, hence both adults and 8-month-old infants can distinguish non-linguistic tones. Future studies should address the issue if tones are perceived linguistically in infants. (a) To disambiguate the first possibility, investigation on both elder infants and younger infants with pure (non-linguistic) tonal stimuli may reveal the answer. It may also help to understand whether infants perceive speech and non-speech tones in a different fashion. (b) Whether or not tone is perceived as intonation in Dutch infants cannot be answered in the current experiment. To disentangle the second possibility, follow up studies should test infants exposed to non-stress timed languages, such as French or Chinese infants. If cross-linguistic comparisons show that the PR pattern is the same across language environments, it is not likely that Dutch infants perceive tone as intonation.

Generally speaking, the findings of the three experiments are in accordance with previous studies

(Capel, 2008; Maye et al., 2002, 2008; Yoshida et al., 2010). In the current experiment, 11-month-old infants' perceptual discrimination is influenced by the frequency distribution of lexical tones in the familiarization phase. This goes with the findings in Maye and Capel's studies in which infants show a similar pattern when hearing distinct distributions of consonants. Moreover, the different perception between infants of 11 months and those of 14 months in the current study matches the distinct perceptual bias between infants of 6–8 months and those of 10 months in the study of Yoshida et al. (2010). After the same duration of exposure to the uni/bimodal, the younger infants are able to catch the distributional pattern while the older infants encounter difficulty. The reason may be that older infants have gathered more distributional experiences outside the lab in the ambient environment so that lab experience has relatively smaller influence. Interestingly, while the same amount of distributional information is not sufficient enough to change older infants' phonemic discrimination, an extended learning period may still make it possible for the infants to "recover" the distributional phonemic learning. To test whether exposure period has an influence on perception, the duration (and hence frequency) of the current experiment should also be doubled in 14-month-old infants. If these infants show the same pattern as those in Yoshida et al. (2010), it is likely that statistical learning is at least one of the mechanisms underlying PR besides phonemic categories per se and potential contextual factors.

In the end, several things are worth discussing. First, infants' ability to perceive non-native speech features is not lost but simply attenuated (Burnham & Mattock, 2007). In each language group tested, there are infants who notice the difference. This is even true to one of the 14-month-old infants in the bimodal condition. Hence, although the underlying mechanism of discrimination in different ages might not be the same, the influence of statistical learning is by no means absolute.

Second, the statistical learning discussed in this paper is merely one of the many types. It is likely that different types of statistical learning, or even some minor changes of the setting of the current type, will lead to a different impact on the results. Interestingly, at least sequential statistical learning is argued to be domain-general[10] (Saffran et al. 1999), whereas PR seems to be linguistically based

---

[10] For instance, infants are able to track musical tonal sequences and find tone-word boundaries through statistical cues (e.g., Saffran, 2003a; Saffran & Griepentrog, 2001; Saffran, Johnson, Aslin, & Newport, 1999), and they can

(Mattock & Burnham, 2006). Suppose a separated mechanism hypothesis is adopted here, it would be interesting to ask how the interaction between domain-specific and domain-general functions works in brain, which function takes a higher priority, or whether a consistent pattern can be observed when two systems collide. Further studies in neuro-linguistics may shed light on these issues.

A third question would be whether tonal PR ends before or after 11 months, and consequently what's the actual effect of statistical learning in tone perception. To answer this question, control groups set up as baselines need to be tested. Infants will be tested on their discrimination without distributional interference in the familiarization phase. By doing this, the result will not only reveal whether 11 months are facilitated or impeded (whether tonal PR ends before or after 11 months), but also show the actual influence of statistical learning in each step of tonal PR.

A fourth question lies in whether tonal language exposure has an influence on foreign tone perception. To find out the answer to this question Limburg (a dialect of Dutch in which lexical tones exist) infants need to be recruited and tested. It would be interesting to compare the Dutch infants and Limburg infants on their performance to tonal perception.

Last but not least, it remains unclear whether tonal PR is dependent also on the tonal contrast tested, although it is likely to be the case based on the current findings. To verify this, more tonal contrasts should be tested within the same age range and same condition. If at some point infants are able to distinguish certain tonal contrast (e.g. T1–T4), but not for some others (e.g. T2–T3), then the answer should be affirmative to this question.

---

learn statistically defined visual patterns (e.g., Fiser & Aslin, 2002; Kirkham, Slemmer, & Johnson, 2002).

Bibliography

Adriaans, F. & Kager, R. (2010). Adding generalization to statistical learning: The induction of phonotactics from continuous speech. *Journal of Memory and Language, 62, 311–331.*

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

Aslin, R.N., & Pisoni, D.B. (1980). Some developmental processes in speech perception. *In G.H. Yeni-Komshian, J.H. Kavanagh, & C.A. Ferguson (Eds.), Child phonology, 2: Perception (pp. 67–96). New York: Academic Press.*

Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P.W., Kennedy, L. J. & Mehler, J. (1988). An investigation of young infants' perceptual representations of speech sounds. *Journal of Experimental Psychology: General, 117(1), 21–33.*

Best, C.T., McRoberts, G.W., & Sithole, N.N. (1988). Examination of the perceptual reorganization for speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14, 345-360.*

Best, C.T., McRoberts, G.W., & Goodell, E., (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of Acoustic Society of America. 109 (2), 775-794.*

Best, C. T. & McRoberts, G. W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Lang. Speech 46, 183–216.*

Boersma, P. & Weenink, D. (2009): Praat: doing phonetics by computer (Version 5.1.05) *[Computer program]. Retrieved May 1, 2009, from http://www.praat.org/*

Bonatti, L.L., Peña, M., Nespor, M. & Mehler, J. (2005). Linguistic constraints on Statistical Computations. *Psychological Science, 16(6), 451-9.*

Bosch, L, Costa, A., & Sebastián-Gallés, N. (2000). First and second language vowel perception in early bilinguals. *European Journal of Cognitive Psychology, 12, 189–221.*

Bosch, L. & Sebastián-Gallés, N. (2003b) Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Lang. Speech 46, 217–24.*

Bosch, L. & Sebastián-Gallés, N. (2005). Developmental changes in the discrimination of vowel contrasts in bilingual infants. *In J. Cohen, K. McAlister, K. Roltad, & J. MacSwan, (Eds.), ISB4: Proceedings of the 4th International Symposium on Bilingualism (pp. 354–363). Somerville, MA: Cascadilla Press.*

Bosch, L. & Sebastián-Gallés, N. (2009). Developmental shift in the discrimination of vowel contrasts in bilingual infants: is the distributional account all there is to it? *Developmental Science. Vol. 12, No. 6, pp. 874-887*

Brown-Schmidt, S. & Canseco-Gonzalez, E. (2004). Who Do You Love, Your Mother or Your Horse? An Event-Related Brain Potential Analysis of Tone Processing in Mandarin Chinese. *J Psycholinguist Res, 33(2):103-35.*

Burnham, D.K. (1986). Developmental loss of speech perception: exposure to and experience with a first language. *Applied Psycholinguistics 7, pp. 207–240.*

Burnham, D., & Francis, E. (1997). The role of linguistic experience in the perception of Thai tones. *In A. S. Abramson (Ed.), Southeast Asian linguistic studies in honour of Vichin Panupong (pp. 29–46). Bangkok: Chulalongkorn University Press.*

Burnham, D. & Mattock, K. (2007). The perception of tones and phones. *Chapter in O.-S Bohn & M. J. Munro (Eds): Language Experience in Second Language Speech Learning: In honor of James Emil Flege. Amsterdam: John Benjamins (Series: Language learning and Language Teaching), Chapter 15, 259-280.*

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

Capel, D.J.H. (2008) Statistical cues affect phonetic discrimination in Dutch infants. *Master thesis, Uil-OTS, Utrecht University*

Clumeck, H.V. (1977). Studies in the acquisition of Mandarin phonology. *Unpublished doctoral dissertation, University of California, Berkeley.*

Clumeck, H. (1980). The acquisition of tone. *In G. H. Yeni- Komshian, J. F. Kavanaugh, & C. A. Ferguson (Eds.), Child phonology: Vol. 1. Production (pp. 257–275). New York: Academic Press.*

Coster,D.C. & Kratochvil, P. (1984). Tone and stress discrimination in normal Peking dialect speech. *In B. Hong (Ed.), New papers in Chinese linguistics (pp. 119 – 132). Canberra: Australian National University Press.*

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1989). Limits on bilingualism. *Nature, 340, 229–230.*

De Casper, A.J. & Fifer, W.P. (1980). Of human bonding: newborns prefer their mothers' voices. *Science, Vol 208, Issue 4448, 1174-1176*

Eimas, P.D. (1974). Auditory and linguistic processing of cues for place of articulation by infants. *Perception & Psychophysics, 16, 513-521*

Eimas, P.D. (1975a). Auditory and phonetic coding of the cues for speech: Discrimination of the [r-l] distinction by young infants. *Perception & Psychophysics, 18, 341-347.*

Eimas, P.D. (1975b). Speech perception in early infancy. *In L.B. Cohen & P. Salapatek (Eds.), Infant perception. New York: Academic Press.*

Eimas, P.D., Siqueland E.R., Jusczyk, P. & Vigorito, J. (1971). Speech Perception in Infants. *Science, Vol. 171, no. 3968, pp. 303 - 306*

Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff- Smith, A., Parisi, D. & Plunkett, K. (1996). Rethinking innateness: A connectionist perspective on development. *Cambridge, ma: The mit Press.*

Elman, JL., Diehl, RL., & Buchwald, S.E. (1977). Perceptual switching in bilinguals. *Journal of the Acoustical Society of America, 62, 971–974.*

Fernald, A. & Kuhl, P.K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, Volume 10, Issue 3,279-293.*

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

Flege, J. E. (2003). Accessing constraints on second-language segmental production and perception. *In A. Meyer & N. Schiller (Eds.), Phonetics and phonology in language comprehension and production, differences and similarities (pp. 319–355). Berlin: Mouton de Gruyter.*

Francis A.L., Ciocca, V., Ma, L. & Fenn, K.. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tonal language speakers. *J Phonetics, 36: 268–294*

Gandour, J., Dzemidzic, M., Wong, D., Lowe, M., Tong, Y., Hsieh, L. & Lurito, J. (2001). Differential neural circuitry underlies the perception of Chinese tone and intonation: an fMRI study. *NeuroImage 13(6): S533. 7th International Conference on Functional Mapping of the Human Brain, June 10-14, Brighton, England.*

Goldsmith, J. (1976). Autosegmental phonology. *PhD diss. MIT, Cambridge MA. Distributed by IULC, published by Garland Press, New York, 1979.*

Goldsmith, J. (1990). Autosegmental and Metrical Phonology. *Blackwell, Oxford.*

Gomez, R.L. & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. Cognition; 70(2): 109-35.

Grosjean, F. (1997). Processing mixed language: Issues, findings and models. *In A. de Groot & J. Kroll (Eds.), Tutorials in bilingualism: Psycholinguistic perspectives. Mahwah, NJ: Erlbaum.*

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism:*

Language and Cognition, 1, 131–149.

Halle P., Chang, Y-C. & Best, C. (2004) Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics, 32, 395-421.*

Harrison, P. (2000). Acquiring the phonology of lexical tone in infancy. *Lingua, Volume 110, Issue 8, August 2000, Pages 581-616*

Houston, D.M., Horn, D.L., Qi, R., Ting, J. & Gao, S. (2007) Assessing speech discrimination in individual infants. *Infancy, 12, 119-145*

Hsieh, L., Gandour, J., Wong, D. & Hutchins, G. (2000). PET investigation of the perception of Chinese tones. *6th International Conference on Functional Mapping of the Human Brain, June 12-16, San Antonio.*

Jusczyk, P.W. (1992). Developing phonological categories from the speech signal. *In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), Phonological Development: Models, Research, Implications. Timonium, MD: York.*

Jusczyk, P.W. (1993). From general to language-specific capacities: The WRAPSA model of how speech perception develops. *Journal of Phonetics [Special issue on phonetic development], 21(1–2), 3–28.*

Jusczyk, P.W. (1997). The discovery of spoken language. *Cambridge,MA: Bradford Books: MIT Press.*

Jusczyk,P.W., Friederici, A.D., Wessels, J.M., Svenkerud,V.Y. & Jusczyk, A.M.(1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory & Language, 32 (3), 402–420.*

Jusczyk, P.W., Luce, P. A. & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory & Language, 33(5), 630–645.*

Jusczyk, P. W., Pisoni, D. B., Walley, A. C., & Murray, J. (1980). Discrimination of the relative onset time of two-component tones by infants. *Journal of the Acoustical Society of America, 67, 262–270.*

Jusczyk, P.W. & Thompson, E.J. (1978). Perception of a phonetic contrast in multisyllabic utterances by two-month-old infants. *Perception & Psychophysics, 23, 105-109*

Kaan, E., Barkley, C.M., Bao M-Z. & Wayland, R. (2008). Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese: An event-related potentials training study. *Neuroscience, 9:53*

Kaplan, E.L. and Kaplan, G.A. (1971). The prelinguistic child. *In: Eliot, J., Editor, Human* *development and cognitive processes, Holt, Rinehart, & Winston, New York, pp. 358–381.*

Klein,D., Zatorre,R., Milner,B. & Zhao, V. (2001). A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *Neuroimage, 13,646 –653.*

Kong, Q-M. (1987). Influence of tones upon vowel duration in Cantonese. *Language and Speech, 30,387 –399.*

Kuhl, P.K. (2000). A new view of language acquisition. *PNAS vol.97, No.22, pp. 11850-11857.*

Kuhl, P.K. (2008). Linking infant speech perception to language acquisition: Phonetic learning predicts language growth. *In Press. McCardle, J. Colombo & L. Freund (Eds.), Infant pathways to language: Methods, models, and research directions (pp. 213-243). Erlbaum: New York.*

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M. & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B, 363, 979-1000.*

Kuhl, P.K., Conboy, B. T., Padden, D., Nelson, T. & Pruitt, J. (2005) Early speech perception and later language development: Implications for the “critical period.” *Language Learning and Development, 1, 237-264*

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, F13–F21.*

Kuhl, P.K., Tsao, F.M., & Liu, H.M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, 100, 9096-9101.*

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

Lee, L. & Nusbaum, H. (1993). Processing interactions between segmental and suprasegmental information in native speakers of English and Mandarin Chinese. *Perception and Psychophysics, 53, 157 –165.*

Li, C.N., & Thompson, S.A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language, 4, 185–199.*

Marcus, G.F., Vijayan, S., Rao, S.B. & Vishton, P.M. (1999). Rule learning in seven-month-old infants. *Science, 283, 77-80.*

Mattock, K. & Burnham, D. (2006). Chinese and English Infants’ Tone Perception: Evidence for

Perceptual Reorganization. *Infancy, 10(3), 24-265*

Mattock, K., Molnar, M., Polka, L. & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition, 106,1367 –1381*

Maye, J. (2000). The Acquisition of Speech Sound Categories on the Basis of Distributional Information. *Unpublished Ph.D. dissertation, University of Arizona, Tucson, AZ, USA.*

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

Maye, J., Weiss, D.J. & Aslin, R.N. (2008) Statistical phonetic learning in infants: facilitation and feature generalization. *Developmental Science, 11:1, pp 122-134*

Mehler, J., Jusczyk, P., Lambertz, G., Halstead, N., Bertoncini, I. & Amiel-Tison, C. (1988). A Precursor of Language Acquisition in Young Infants. *Cognition, 29, 143-78*

Mennen, I., Levelt, C. & Gerrits, E. (2006). Acquisition of Dutch phonology: an overview. *QMUC Speech Science Research Centre Working Papers, WP-10.3*

Miller, J.L. & Eimas, P.D. (1979). Organization in infant speech perception. *Canadian Journal of Psychology, 33,* 353-367.

Morse, P.A. (1972). The Discrimination of Seech and Nonspeech Stimuli in Early Infancy. *Journal of Experimental Child Psychology 14, 477-492.*

Nazzi, T., Floccia, C. & Bertoncini, J. (1998). Discrimination of pitch contours by neonates. *Infant Behavior and Development, 21, 779 –784.*

Nelson, D.G.K., Jusczyk, P.W., Mandel, D.R., Myers, J., Turk, A. & Gerken, L.A. (1995). The Head-Turn Preference Procedure for Testing Auditory Perception. *Infant Behavior and Development 18, 111-116.*

Nespor, M., Peña, M. & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio, 2, 221–247.*

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

Pegg, J.E., & Werker, J.F. (1997). Adult and infant perception of two English phones. *Journal of the Acoustical Society of America, 102, 3742– 3753.*

Polka, L., Colantonio, C. & Sundara, M. (2001). A cross-language comparison of ⁄d⁄–⁄ð⁄ perception: Evidence for a new developmental pattern.

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

Repp, B.H. & Lin, H.B. (1990). Integration of segmental and tonal information in speech perception: A cross-linguistic study. *J. Phonetics 18, 481–495.*

Sabourin, L., Bosch, L., Sebastián-Gallés, N., & Werker, J. F. (2006). Perceiving vowels in a tight vowel space: Evidence from monolingual infants. *Manuscript submitted for publication.*

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

Saffran, J.R., Johnson, E.K., Aslin, R.N. & Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, Volume 70, Issue 1, 1 February 1999, Pages 27-52*

Streeter, L.A. (1976). Kikuyu labial and apical stop discrimination. *Journal of Phonetics, 4, 43-49.*

Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: bilingual infants keep pace with their monolingual peers. *Cognition, 108, 232–242.*

Trehub, S.E. (1976). The Discrimination of Foreign Speech Contrasts by Infants and Adults. *Child Development, 47, 466-472.*

Tse, J. K. (1978). Tone acquisition in Cantonese: A longitudinal study. *Journal of Child Language, 5, 191–204.*

Tuaycharoen, P. (1977). The phonetic and honological development of a Thai baby: From early communicative interaction to speech. *Unpublished doctoral dissertation, University of London.*

Van Lancker, D., & Fromkin, V. (1973). Hemispheric specialization for pitch and “tone”: Evidence from Thai. *Journal of Phonetics, 1, 101-109.*

Veenker, T.J.G (1998). FEP, A tool for designing and running computerized experiments. *Utrecht Institute of Linguistics – OTS, Utrecht University, Utrecht, The Netherlands.*

Veenker, T.J.G. (2007). *[computer program]. http://www.let.uu.nl/~Theo.Veenker/personal/ZEP.*

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

Wang, W.S.Y. (1973). The Chinese language. *Scientific American, 228, 51-60.*

Wang, Y., Behne, D.M., Jongman, A. & Sereno, J.A. (2004). The Role of Linguistic Experience in the Hemispheric Processing of Lexical Tone. *Applied Psycholinguistics, vol. 25, no. 3, pp. 449-466.*

Wang,Y., Jongman,A. & Sereno,J.A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language, 78,332 －348.*

Werker, J. F. (1994). Cross-language speech perception: Developmental change does not involve loss. *In J. Goodman & H. Nusbaum (Eds.), The development of speech perception: The transition from speech sounds to spoken words (pp. 112-149). Cambridge, MA: MIT Press*

Werker, J.F. & Byers-Heinlein, K. (2008) Bilingualism in infancy: First steps in perception and comprehension of language. *Trends in Cognitive Sciences, 12(4), 144-151.*

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. & Logan, J.S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics, 37, 35 – 44*

Werker, J.F. & Pegg, J.E. (1992). Infant speech perception and phonological acquisition. *In C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), Phonological development: Models, research, and implications (pp. 285-311). York Publishing Company.*

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

Werker, J.F. & Tees, R.C. (2002). Cross-language speech perception: Evidence for perceptual reorganization in the first year of life. *Infant Behavior and Development, 25, 121–133.*

Werker, J.F. & Tees, R.C. (2005). Speech Perception as a Window for Understanding Plasticity and Commitment in Language Systems of the Brain. *Wiley Periodicals, Inc. Dev Psychobiol 46: 233–251*

Whalen, D. & Xu, Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica, 49, 25 －47.*

Wong, P., Schwartz, R.G. & Jenkins, J.J. (2005). Perception and Production of Lexical Tones by 3-Year-Old, Mandarin-Speaking Children. *Journal of Speech, Language, and Hearing Research, Vol. 48, 1065–1079*

Wu, X.H., Lin, H. (2008) Perception of Mandarin Tones by Mandarin and English Listeners. *Journal of Chinese Language and Computing, 18(4), 175-187*

Yoshida, K., Pons, F., Cady, J.C. & Werker, J.F. (2006). Distributional learning and attention in phonological development. *Poster presented at the International Conference on Infant Studies, Kyoto, Japan.*

Yoshida, K.A., Pons, F., Maye, J. & Werker, J.F. (2010). Distributional phonetic learning at 10 months of age. *Infancy, 15(4), 420-433.*

Zhu, H., & Dodd, B. (2000). The phonological acquisition of Putonghua (modern standard Chinese). *Journal of Child Language, 27, 3–24.*

Appendix I – Adult experiment

Identification and discrimination of a tonal contrast in Mandarin Chinese by Chinese and Dutch adult listeners

1 Introduction

In tonal languages, different tones distinguish lexical or grammatical meanings (Wang, 1973). Four overt tones appear in Mandarin Chinese, marked in numbers from 1 to 4. Tone 1 (T1) is a high-level tone, while tone 4 (T4) is high-falling. In this language, each word is mono-syllabic and bears one tone. Same syllables with different tones form lexically contrastive minimal pairs (e.g. /ma1/ means mother while /ma3/ means horse; the number after the syllable represents the tone). In phonology, tones can be seen as phonetic distinctions attached to the syllable at a prosodic level. Their properties include fundamental frequency (F0), amplitude (intensity) and duration (Coster and Kratochvil, 1984; Halle et al., 2004; Kong, 1987; Whalen and Xu, 1992).

Previous ERP, PET and fMRI studies suggest that lexical tones are perceived in different domains in the brain by native and non-native tonal-language listeners. They seem to be perceived as linguistic information among native listeners and are processed in the left hemisphere, just as other speech segments (e.g. vowels) (Schmidt and Gonzalez, 2004; Gandour et al., 2001; Hsieh et al., 2000; Kaan et al., 2008; Klein et al., 2001; Wang et al., 2001), while by non-native listeners of a tonal language[11], they seem to be perceived as non-linguistic melodic information (Van Lancker and Fromkin, 1973; Wang et al., 2004). Native speakers of a non-tonal language also show a degree of sensitivity to supra-segmental information such as stress and intonation (Repp and Lin, 1990; Lee and Nusbaum, 1993). Different native languages (L1) seem to evoke different sensitivities to the prosodic variation of a second language (L2) (Soto-Faraco et al., 2001). Dutch speakers may have a high sensitivity to syllable-level prosodic variation, since not only is Dutch a stress language involving rich prosodic information, but it also has several lexically contrastive minimal pairs marked by stress placement (e.g. canon and kanon). The potential L1 influence on L2 perception on a supra-segmental level is particularly intriguing.

1.1 Categorical perception (CP) of tones

Recent research suggests that tones are categorically perceived by native speakers of a tonal language and that tonal variation is used in a phonologically contrastive way at the lexical level. Priming studies show that tonal categories function to distinguish between segmentally identical words in native listeners of a tonal language (Lee, 2007). Meanwhile, native speakers of a non-tonal language process prosodic variation at the sentence level in a loose fashion (c.f. intonation) (Halle et al., 2004; Xu et al., 2006; Wang, 1976; Wu and Lin, 2008). Francis et al. (2008) provide a category-based interpretation of the perceptual learning of Cantonese lexical tones by Chinese (native tonal language) and English (non-native tonal language) listeners and attribute their performance not solely to subjects' L1 tonal categories, but also to cross-language mappings

---

[11] But see Luo et al. (2006) for counter-examples.

including prosodic knowledge of both L1 and the target language. Francis et al. (2003) show CP traces for tonal language listeners in lexical tone identification tasks. However, in discrimination tasks, the trace seems to evaporate. Similar results show up in Abramson's (1979) study with Thai tones, in which the author argues that no CP in Thai tones can be found in discrimination tasks. Note that these findings seem to contradict each other in the claim of CP of tones. This not only calls for more experiments on other tone (non-tone) language speakers, but also for a better integration and interpretation of tasks used in CP experiments.

To take this step, it is useful to bear two points in mind. Firstly, native speakers of a tonal language must ignore within-category tonal variations in order to categorize f0 contours into tones accurately and efficiently. They should be less sensitive than speakers of a non-tonal language in discriminating small f0 contour variations (Stagray and Downs, 1993). Secondly, in the general CP literature, there is no clear definition of category prototypes, nor is there agreement on the definition of a categorical boundary. Hence, the CP criteria should be elaborated in discussing CP of lexical tones. This point will be discussed later in the section on the degree of CP.

1.2 Literature review

Two recent papers provide information about a CP effect in tonal perception of Mandarin Chinese[12]. Halle et al. (2004) set up 8-step continua for three sounds /pa/, /pi/ and /kwo/ in three tonal contrasts, respectively: T1-T2; T2-T4; and T3-T4. Three experiments were designed to test the potential CP effect among Taiwanese and French subjects. In the first experiment, 15 Taiwanese were tested on a forced choice identification task and a forced choice AXB discrimination task. Results show that tone perception by Taiwanese listeners is relatively categorical. In the second experiment, 14 Taiwanese and 14 French students participated in an AXB identification task[13] on single syllables. Results show that Taiwanese and French subjects differ qualitatively in categorizing tones: the French subjects' identification performance reflects psychophysical rather than linguistic sensitivity towards lexical tones, while Taiwanese subjects follow the phonological value assigned to tones in Mandarin Chinese. In the third experiment, an AXB forced choice discrimination task which was similar to the first experiment was adopted. The results of 14 French subjects again reveal that their perception is not categorical, but is psychophysically based. In brief, in Halle et al.'s study only native Mandarin subjects show CP towards lexical tones.

Using PRAAT software, Xu et al. (2006) created 7-step continua of speech (from level to rising tone) and nonspeech stimuli. These stimuli were used in a forced choice identification task and an AX discrimination task. Data from 15 Chinese and 15 English listeners reveal a CP effect for Chinese but not for English listeners. Xu et al. argue against auditory mechanisms and favour a memory factor explanation to account for their finding of the same identification boundaries near the middle of the continuum regardless of language background. This raises interesting questions for the current study, such as whether the categorical boundary locations of the Chinese and Dutch groups will be close to each other.

---

[12] In this paper Mandarin Chinese may also be called Mandarin or Chinese.

[13] The difference between an AXB identification task and an AXB discrimination task is that in the identification task, A and B correspond to two endpoints while X varies from one endpoint to the other along the 8-step continuum, whereas in the discrimination task X is equal to the sound A or B.

In summary, in addition to Thai, Vietnamese and Cantonese (Abramson, 1979; Schwanhaeusser, 2008; Francis et al. 2003), the CP effect of Mandarin Chinese has also been tested in English, French and as Chinese listeners. Identification tasks result in a CP pattern for Chinese listeners while discrimination tasks seem to provide less evidence. Although neither English nor French listeners show CP in lexical tones, the explanations vary from psychophysical to memory factors.

No previous literature has addressed how Dutch listeners perceive lexical tones. Being a stress-based language, Dutch is quite different from French or Chinese (both syllable-based) in linguistic rhythm. It would be very interesting to see how Dutch subjects would perform the tasks. Moreover, few papers have overtly discussed the relationship between identification and discrimination tasks, which may result in a stronger interpretation of the CP effect.

1.3 Aims of the study

The aims of this study were to answer the following questions: (1) Do native Chinese subjects perceive the native lexical tonal contrast between T1 and T4 categorically? What will the boundary location be? (2) To what extent will native Dutch subjects identify and discriminate lexical tones compared to Chinese subjects? (3) In what way will native Dutch speakers perceive the differences, linguistically (categorically) or non-linguistically?

Since a cross-linguistic difference seems to exist between native speakers of tonal and non-tonal languages in the perception of tones, it is predicted that Mandarin Chinese subjects will show CP when perceiving lexical tones and will treat them as contrastive linguistic categories, while Dutch subjects will not reveal the same perception pattern.

2 Method

2.1 Participants

In total, 44 adult subjects participated in the study. They were 23 native Mandarin Chinese and 21 native Dutch subjects currently living in Utrecht, the Netherlands. All subjects reported normal hearing and speaking. Three native Chinese subjects were dropped from the study because they either did not understand the instructions or were not naive as to the purpose of the experiment. One native Dutch subject was dropped from the study due to her previous linguistic experience in Mandarin Chinese. No other Dutch participants had ever been exposed to Mandarin Chinese or any other tonal languages; neither had any of them ever travelled to China. After the filtering, 20 native Chinese and 20 native Dutch speakers' (mean age 25 years) performance was examined in this study, with 6 males and 14 females per language group.

2.2 Stimuli

The speech stimuli used for the experiments were from the /ta1/-/ta4/ tonal contrast. In Chinese, syllable /ta/ with different tones has distinct meanings. (e.g. /ta1/ as "put up" and /ta4/ as "big") Both sounds were produced four times by the same female speaker to add within-speaker acoustic variation, which allowed the stimuli to resemble natural speech during testing. After the recording, these sounds were resynthesized in PRAAT software to generate four continua of the same contrast, marked as continua A, B, C and D respectively. The F0-range between the two speech sounds /ta1/ and /ta4/ was divided into 7 equal distances by specifying 7 F0 values at 4 points in time, since this caused a more natural synthesized speech sound compared to linearly falling contours (with only two points in time). Eight stimuli were obtained per continuum. The other acoustic properties (e.g. intensity) of the tonal stimuli were left as they were. The average duration of the stimuli was 0.412 s (ranging from 0.376s to 0.448 s). In total, 4 * 8 = 32 stimuli were generated to make the four continua of /ta1/-/ta4/. Five native speakers of Mandarin Chinese listened to them and agreed that all 32 stimuli from the continua sounded like normal speech. For all tasks, the stimuli were randomized but were always from the same continuum within each trial.

2.3 Procedure

All subjects participated in a set sequence of four tests: a forced choice identification task, an AX discrimination task, a free choice identification task and an AXB discrimination task. The total duration of these experiments added up to 25-30 minutes, depending on the individual speed of the subjects. Subjects could take a rest in between the tasks. Proper instructions were provided before the tasks, including a very brief introduction to tonal language. Considering that it would be unfair for Dutch listeners to choose between Mandarin tones, auditory examples were added in identification tasks (see section 3.1).

The experiments were conducted in the Phonetics Laboratory of Utrecht University. The scripts were all created from the Veenker (1998) FEP programming system. Each subject was tested individually. They were seated in front of a monitor in a sound-attenuated booth and stimuli were presented at a sound level adequate to the listeners.

3 Experiment

3.1 Task description

Two identification tasks (tasks 1 and 3) and two discrimination tasks (tasks 2 and 4) were set up. Response data were collected for all tasks and reaction time (RT) data were collected for discrimination tasks. In all tasks subjects were aware of the fact that they were listening to tonal linguistic stimuli.

Choosing these four tasks was not an arbitrary decision. An absolute identification task revealed more information of the subjects that might be missing in a forced choice task. Besides that, due to the large variation in findings in the literature, it is better to take a comprehensive look from both the identification and the discrimination angle. Since the performance of subjects on two tasks inherently tapping into the same function should be similar, it is useful to do 2 tasks, one of which serves as a replication. Moreover, a 1-identification X 1-discrimination combination of experiments only leads to one comparison, whereas a 2 X 2 design results in a more powerful 4 comparisons in total, which is more likely to reveal the true relationship between the two functions.

3.1.1 Task 1 – forced choice identification

In this task, subjects heard one stimulus per trial. Before the actual test, acoustic examples A1 (continuum A, stimulus 1, the flattest sound), B8 (the most falling sound) and C5 (a sound close to the middle) were played once each, in order for subjects to understand the difference between T1 (A1) and T4 (A8). Native Chinese subjects were required to choose between “tone1” and “tone4”, shown as two buttons on the screen, whereas native Dutch subjects were asked to choose between “flat” and “falling”[14] sounds after hearing each stimulus. Subjects were told to respond to the stimulus as accurately and quickly as possible by clicking one of the two buttons on the screen. After each click, the next trial was presented. All 32 stimuli were played twice in random order, adding up to 64 trials for each subject. In this task, each stimulus was marked as a step and therefore there were 8 steps in total along the continuum.

3.1.2 Task 2 – AX discrimination

In this task, subjects heard pairs of stimuli along one continuum and were required to make a forced choice on whether the two stimuli they heard were the same or different[15]. Before the actual test, examples A1-A1 pair (same) and B1-B8 pair (different) were played as illustrations. All subjects were asked to respond to the pairs as accurately and quickly as possible by clicking one of the two buttons, labeled “same” and “different”, on the screen. After each click, the next trial was presented. Subjects’ response and RT to 0-step trials (1-1, 2-2, etc.) and 2-step trials (1-3, 2-4, 3-5, 4-6, 5-7 and 6-8) for each continuum were recorded, adding up to 128 trials for each subject. To prevent experiment-induced bias, 48 extra trials with different pairs (the stimuli in these trials were more than 3 steps apart, e.g. 2-6 pair) were used as filler trials.

3.1.3 Task 3 – free choice identification

In this task, subjects heard one stimulus per trial and they were asked to score the stimulus they heard from 1 to 8, with score 1 resembling T1 (being the flattest sound) and score 8 representing T4 (being the most falling sound). Before the actual test, acoustic examples A1, B8 and C5 were played once each, in order for subjects to get familiar with the scoring task. In an ideal case, for example, if a subject heard stimulus A4, he or she would score 4 as the standard answer without bias. Subjects were told to respond to the stimulus as accurately and quickly as possible by clicking one of the eight buttons (from 1 to 8) on the screen. After each click, the next trial was presented. All 32 stimuli were played twice in random order, adding up to 64 trials for each subject. In this task, each stimulus was marked as a step and therefore there were 8 steps in total along the continuum.

3.1.4 Task 4 – AXB discrimination

In this task, subjects heard three stimuli per trial along one continuum and were required to make a forced choice as to whether the second stimulus (X) sounded the same as or closer to the first stimulus (A) or the third one (B). The AXB trials had four possible combinations (AAB, ABB, BAA and BBA). Before the actual test, acoustic examples A1-A1-A8 triplet (X=A) and A1-A8-A8 triplet (X=B) were played as illustrations. All subjects were asked to respond to the triplets as accurately and quickly as possible by clicking one of the two buttons, labeled “first” and “third”, on the screen. After each click, the next trial was presented. (The response and RT of 2-step (AB) trials (1-3, 2-4,

---

[14] T1 is acoustically flat while T4 is a falling sound. Therefore, the different button names match the acoustic properties of the tones. This setting tends to be fair for Dutch listeners since they have no knowledge about what T1 and T4 represent. The difference in choice may have a potential impact on the performance of Dutch subjects, (see section 3.2).

[15] In an AX discrimination task, each trial consists of 2 stimuli (A and X) and subjects need to discriminate them.

3-5, 4-6, 5-7 and 6-8) for each continuum were recorded, adding up to 96 trials for each subject. To prevent experiment-induced bias, 48 extra trials with distinct pairs (the stimuli in these trials were more than 3 steps apart, e.g. 2-6 pair) were used as filler trials. All trials were played in random order..

3.2 Results and discussion

A univariate ANOVA with subject, source, step and language as factors shows that the source factor (4 continua generated by the same person) does not yield a significant difference across tasks, while the subject factor always yields significant differences due to the high variation of subjects' performance. The source factor is thus considered not informative to the current study and is aggregated in later analysis..

3.2.1 Task 1 – forced choice identification

In general, repeated measures ANOVA with language and step as fixed factors shows a significant difference between the two language groups ($p < 0.001$). As FIG 1 shows, Chinese subjects' performance reveals a high consistency, with quite a strong tendency to separate the two tonal targets at both two ends of the continuum. This is especially obvious in the right-hand part of the continuum. The category boundary seems to be located between stimuli 2 and 4. On the other hand, no such pattern can be observed among Dutch listeners, whose performance shows great variability in almost each step along the continuum. There is a difference between step 1 and the rest of the steps: Dutch subjects are able to separate a flat sound from a falling one (steps 1 and 2), while many Chinese subjects seem to consider these 2 steps as being within one category. However, even this difference is not perceived by all Dutch subjects.

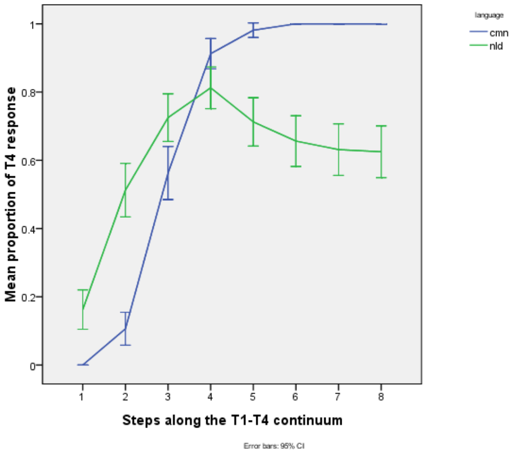


FIG 1 One-step identification curves for Chinese and Dutch subjects

PROBIT analysis (Best and Strange, 1992) was used to calculate category boundary locations and slopes of the Gaussian distribution functions that fitted the identification data. (Halle et al., 2004; Wu and Lin, 2008) In this analysis, the 3 points (out of 8 along the continuum) closest to the 50% mean response value were taken out to fit into an Ogive curve (Gaussian function). These three points were usually in the steepest part of the original curve. The intercept, namely the estimated 50% response position, and the slope, namely the estimated steepness value by 1/SD of the Gaussian function fitted to the data, were generated from the PROBIT program. Generally the intercept indicates the categorical boundary location, while the slope indicates how sharp this boundary is.

However, from FIG 1 it is fairly clear that the curve of the Dutch group does not yield a categorical pattern in the first place and therefore should not be represented by a Gaussian function (nor by a linear equation). Therefore, fitting the results of the Dutch group with a Gaussian curve (or a straight line) would not be meaningful. The intercepts and slopes are summarized in Table 1, with Dutch results just for reference.

Task 1	Intercept		Slope	
	Chinese	Dutch	Chinese	Dutch
/ta1/-/ta4/	2.927	4.117	1.798	0.353

Table 1 Intercept and slope for Chinese and Dutch subjects' forced choice identification data

As expected, the intercept differs significantly between groups: F (1, 22) = 22.61, p = 0.015. The boundary for the Chinese group (the 50% point) is lower than step 3, much to the left of the Dutch group, which is in the middle of the continuum (near stimulus 4). The slope is significantly sharper for the Chinese group: F (1, 38) = 2.8, p < 0.001. The results of the Chinese group resemble a CP function more than those of the Dutch group, which are not categorical.

3.2.2 Task 2 – AX discrimination

Mean accuracy results are pooled in FIG 2 using a repeated measures ANOVA with language and step as factors. The performance of both Chinese and Dutch groups is quite high at the left end (T1), and falls sharply in the middle and at the right end (T4) of the continuum, reaching no significant difference between the two language groups (p = 0.682). For both groups, step 1-3 is the most salient contrast pair (Correctness: Chinese: 93%; Dutch: 100%), while step 6-8 is the most difficult contrast pair for both groups (Correctness: Chinese: 7%; Dutch: 17%). Note that CP predicts that the more accurate the responses are, the more likely it is that the stimuli do not fall into one category.

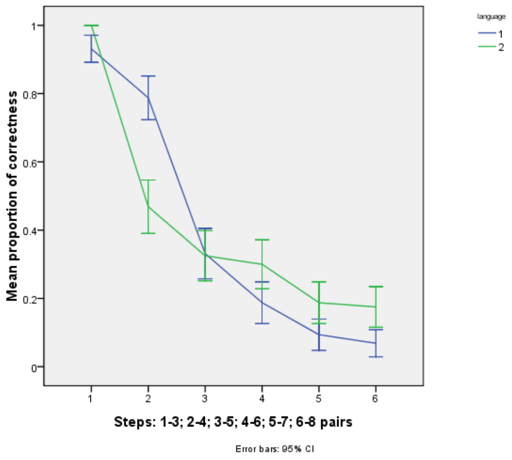


FIG 2 Two-step discrimination curves for Chinese and Dutch subjects

Within the Chinese group, the correct response to steps 1, 2 (pairs 1-3 and 2-4) is significantly higher than all other steps ($p < 0.01$). None of the adjacent steps 3 to 6 are significantly different from their neighbours. In general, Chinese subjects are very poor at distinguishing steps in the middle and on the right of the continuum. From step 4, the accuracy rate drops to less than 20% and becomes even smaller than 10% from step 5 onwards. The curve of the Dutch subjects resembles the Chinese group in this discrimination task, which is not unexpected. Previous literature shows that in such discrimination tasks non-native listeners perform better in perceiving the speech distinction than in identification tasks. They are not completely "deaf" to tonal variation. (Halle et al., 2004; Wu and Lin, 2008).

Comparing the results between the two language groups through univariate ANOVA with the factors of step and language, it can be observed that Dutch subjects seem to outperform Chinese listeners in each of the other 5 steps, except for step 2. It has to be noted that, although Dutch subjects do not perceive the lexical tones (T1 and T4) in a categorical fashion, they are somewhat more sensitive in distinguishing the subtle difference in tonal stimuli. On the other hand, Chinese listeners seem to be "limited" by their L1 tonal category and cannot tell the difference, for instance, in the middle and latter falling part (T4) along the continuum.

The RT results of six 2-step pairs are pooled in FIG 3 with language and step as factors. The general group difference is significant ($p = 0.004$), with a higher RT by Chinese subjects on average. A peak at step 3 can be observed for the Chinese group, while the Dutch group tends to be level on average, except for step 1. Note that the more RT is required, the more difficult the discrimination is and the more likely that the stimuli of a pair are within one category.

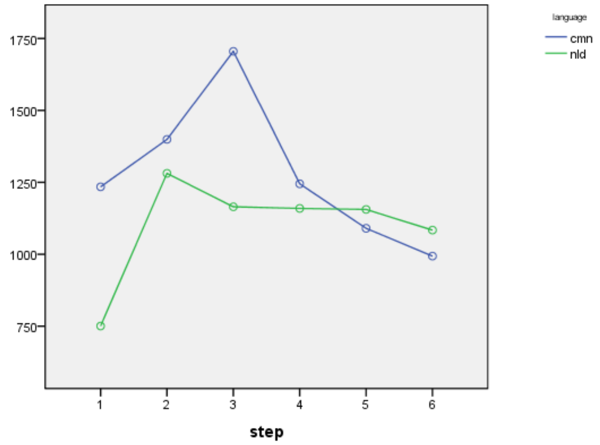


FIG 3 Two-step discrimination curves for Chinese and Dutch subjects in RT

Horizontal step: step 1-3; 2-4; 3-5; 4-6; 5-7; 6-8 pairs Vertical: mean response time (ms)

Within the Chinese group, the peak at step 3 is significantly different from all other steps except for step 2. The other pairs do not differ significantly. RTs at steps 4, 5 and 6 are moderate. Considering the low accuracy rate in these steps, it seems that Chinese subjects do not consider them difficult to judge (although they are actually wrong) and categorize them as within the same category quickly. Step 3 seems to be the hardest for the Chinese listeners, no matter how accurate their decisions are.

The RT peaks in steps 2 and 3 reveals a potential category boundary: the falling sounds within the T4 category are quite similar to Chinese listeners and hence they need more time to make decisions; the rest of the steps are quite obvious. For the Dutch group, only the first step is significantly different from the other steps, whereas the rest of the steps do not yield significant differences ($p > 0.999$). This is a clear illustration of psychophysically based RT. As has been mentioned, the acoustic difference between stimuli 1 and 3 is salient. Hence pair 1-3 does not require much processing time compared to the other steps. RTs of the remaining 5 steps could be considered as equal processing time by listeners, showing no CP trace. In brief, Dutch listeners seem to be quicker and more accurate in this discrimination task.

3.2.3 Task 3 – free choice identification

A comparison is made of the correct response in the original data. As shown in FIG 4, Chinese subjects are quite accurate along both edges of the continuum but not in the middle where they probably categorize the stimuli into either edge. Dutch subjects' response for stimulus 1 (the flat sound) is above chance; however, their accuracy rates radically drop for every other place along the continuum. This favours the claim that stimuli are more categorically perceived by Chinese than by Dutch listeners.

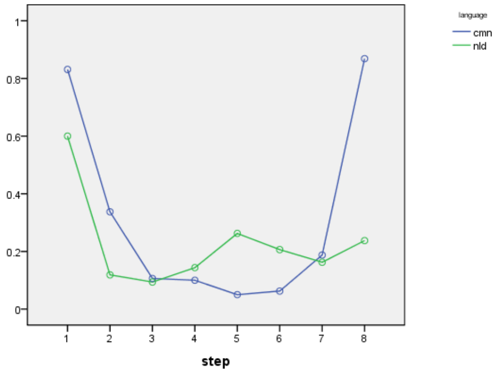


FIG 4 Correct response proportions for Chinese and Dutch subjects

Horizontal: 8 stimuli along the T1-T4 continuum Vertical: accuracy proportion

The specific stepwise trends obtained from subjects' choices strengthen this idea. Taking both groups' response in steps 5, 6 and 7 as examples, Chinese and Dutch listeners perform quite differently when hearing these three stimuli along the continuum. (Table 2)

Listen to	Step 5		Step 6		Step 7
Response	Step7	Step8	Step7	Step8	Step8
Chinese	0.369 (0.03)	0.344 (0.03)	0.300 (0.03)	0.619 (0.03)	0.775 (0.03)
Dutch	0.137 (0.03)	0.056 (0.03)	0.162 (0.03)	0.075 (0.03)	0.094 (0.03)
p value	$p < 0.001$	$p < 0.001$	$p = 0.003$	$p < 0.001$	$p < 0.001$

Table 2 Mean proportion (standard error) of subjects' actual response in steps 5~7

In Table 2, the actual responses of Chinese subjects to the stimuli they hear are much to the higher end of the continuum, 71.3% , 91.9% and 77.5% respectively for steps 5, 6 and 7, which is the main reason for their low accuracy in these steps in the task (FIG 4). It can also be observed that, although

Dutch listeners perform poorly in these steps, they do not make the same kind of mistakes as Chinese listeners. Univariate ANOVA shows that the responses of the two groups are significantly different in all cases of the example.

In brief, although both the Chinese and the Dutch groups are low in accuracy, especially in the middle steps along the continuum, their responses stem from different causes: Chinese subjects are biased by their inner tonal category and seem to prefer choices closer to the "protocol" of the category, while Dutch subjects, without previous linguistic experience of tones, simply fail to distinguish the complex sounds and choose randomly.

The original absolute results have been reduced from eight to two categories, in order to fit a T1-T4 response comparison. In brief, the results of the reinterpreted data of task 3 show consistency with task 1. With step and subject as fixed factors, repeated measures ANOVA shows a significant difference between the two language groups ($p < 0.001$). Similarly to task 1, a strong categorical shift pattern can be observed for the Chinese subjects while a large variation is found among Dutch listeners in FIG 5.

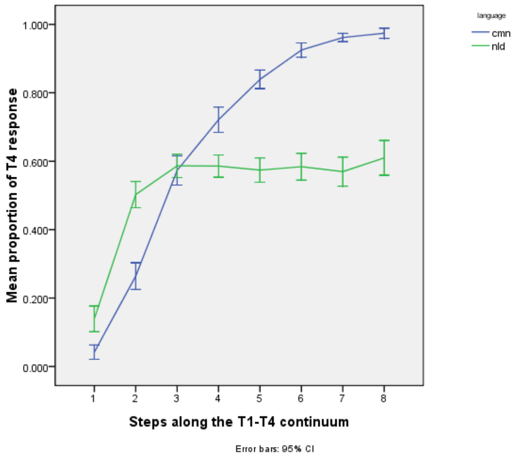


FIG 5 One-step identification curves for Chinese and Dutch subjects in T4 response

Chinese subjects' performances seem to reveal two categories for the two tonal targets at the two ends of the continuum, while Dutch subjects' performance varies at each step, with mean responses moderately above 50% regardless of the step, except for step 1 (the only flat sound). Note that even the flat sound (stimulus 1) is not completely recognized across Dutch subjects. As expected, they show no trace of categorically perceiving the sounds along the continuum.

Once again, data of Chinese subjects fit a Gaussian distribution while data of the Dutch group does not. Therefore, fitting the results of the Dutch group with a Gaussian distribution (or a straight line) would not be meaningful. The identification data for each subject are fitted to the Gaussian curve analyses (same as task 1) to obtain individual and mean intercepts (categorical boundary locations) and slopes (potential degree of CP). The intercepts and slopes are summarized in Table 3, with Dutch results just for reference.

Task 3	Intercept		Slope	
	Chinese	Dutch	Chinese	Dutch
/ta1/-/ta4/	2.86	4.189	0.834	0.253

Table 3 Intercept and slope for Chinese and Dutch subjects' free choice identification data

The intercept values of both groups resemble those of task 1. The intercept again differs significantly between Chinese and Dutch subjects, F (1, 23) = 32.64, p = 0.007. The boundary location (the intercept) result matches not only the identification but also the discrimination tasks. The slopes are weaker for both groups compared to task 1, yet they are still significantly steeper for the Chinese group: F (1, 19) = 9.22, p = 0.012. The picture and the analysis both show that CP is more likely to occur in the Chinese group than in the Dutch group.

3.2.4 Task 4 – AXB discrimination

The data of the two language groups in AXB show much better performance in response accuracy than the AX discrimination task. In fact, for both groups the mean accuracy rate in all steps is above 60% (FIG 6), revealing the highest performance among all tasks. One possible explanation for this could be the learning effect: since the AXB task is the last one of the four tasks using the same stimuli, subjects have been familiarized a lot with the stimuli and pay more attention to the small acoustic differences. Alternatively, it could be that the AXB discrimination task is easier in general, since two comparisons can be made in each trial, adding to the chance of a correct response. At first glance, it seems that both groups display a falling, almost straight line – one falls more steeply than the other. Repeated measures ANOVA with step and language as factors reports a significant difference between the two groups (p = 0.001) and step 1 remains the peak in correctness for both groups.

Chinese subjects' performance near the left edge (steps 1, 2) is quite high (97.5% and 96.2%) and indistinguishable from each other (p > 0.999), while near the right edge (4, 5, 6), their performance is relatively low (66.7% on average), with no difference either (p > 0.999). Repeated measures ANOVA shows that steps 1 and 2 differ significantly from steps 4, 5 and 6, while step 3 stands in the middle of this steep drop in performance and differs significantly from all the other steps. Meanwhile, Dutch listeners perform quite differently from their Chinese peers: although step 1 is significantly different from all other steps (p < 0.001), the differences between all other adjacent steps (e.g. steps 2 and 3; 3 and 4, etc.) are not significant.

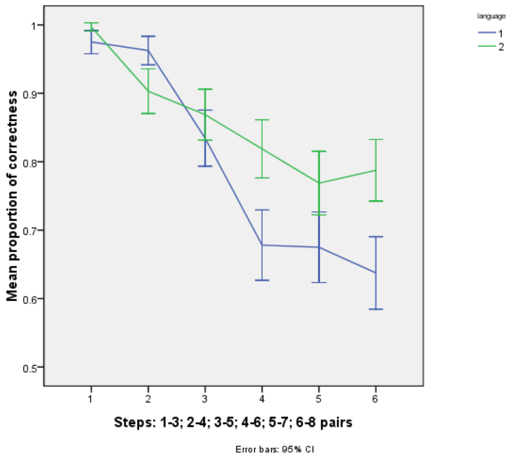


FIG 6 two-step discrimination curves for Chinese and Dutch subjects

Step-wise comparisons between the two language groups through univariate ANOVA reveal very high homogeneity with the results from the AX discrimination task: Dutch subjects outperform Chinese participants in each of the other 5 steps except for step 2. It is likely that whereas Chinese listeners fail to distinguish stimulus pairs that fall within the same category, Dutch listeners, having no linguistic category in lexical tones in their native language, feel more at ease finding the small differences in prosody.

The reliability of RT should be considered low in an AXB discrimination task compared to the other tasks, since some subjects may push the button immediately after the first two sounds and before the third one if they are confident enough to make the decision. Briefly, the overall group comparison reveals an intriguing significant difference ($p = 0.001$) in that Dutch listeners respond faster than Chinese listeners in all steps. It is interesting to see that native speakers of a target tonal language do NOT necessarily perform better than those who have no knowledge of the target language in judging lexical tones.

3.3 Cross-task CP Index analysis

CP *is* the relationship between identification and discrimination tasks: the true identity of CP lies not only in how one categorizes the sounds, but also in the fact that one cannot tell the difference of the sounds within the same category (Gandour, 1978). If subjects do perceive stimuli in a categorical fashion, their performance in identification tasks should determine their discrimination task performance. In principle, the stronger the correlation between the identification and discrimination functions, the higher the degree of CP. In other words, the better identification performance predicts discrimination performance, the more CP there is. The degree of CP can be calculated using the “categorical-perception index” (Van Hessen and Schouten, 1999):

$$CP = [r / (1 + 2 * | p(obt) - p(pred) |)] * 100$$

In this formula, CP refers to the degree of categorical perception. By definition, the index value refers to how categorically the target is perceived. Its value may vary from 0 to 100. The numerator r represents the correlation coefficient between an identification function and a discrimination function. The higher the correlation, the larger the degree of CP. The denominator includes "p(obt) – p(pred)", which represents the mean differences between the discrimination function 'p(obtained)' and the identification function 'p(predicted)' across data. The smaller the difference, the larger the degree of CP.

The discrimination function was obtained directly from the discrimination tasks. It is the averaged accuracy rates (hence discrimination ability) across pairs[16] for each subject. The identification function p(pred) cannot be obtained directly from raw data. It needs to be calculated from the identification differences within the relevant pairs of stimuli. The calculation formula used is called the "Haskins prediction" (Cutting and Rosner, 1976; Macmillan et al. 1977; Pollack and Pisoni, 1971):

$$P(c) = 0.5[1 + (p_1 - p_2)^2]$$

In this formula, P(c) equals p(pred). It refers to the probability of a correct discrimination. P_1 is the probability of assigning stimulus A to a particular category, while p_2 is the probability of assigning stimulus B to the same category. Predicted probabilities can thus be obtained from identification data. The degree of CP can be observed in Table 3.

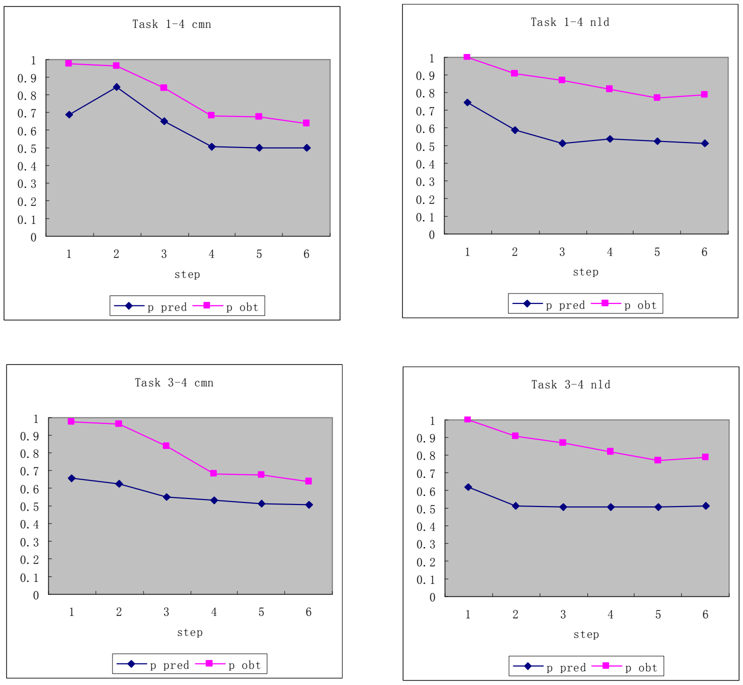
Task combo		CP index	
Identification	Discrimination	Chinese	Dutch
task 1	task 2	41.144	35.974
task 1	task 4	44.382	20.212
task 3	task 2	58.641	52.409
task 3	task 4	48.134	31.482
Mean		48.075	35.019

Table 3 Categorical perception (CP) index for four tasks
(CP index range: 0 – 100)

Regardless of the task combination between the identification and discrimination functions, the degree of CP of Chinese subjects is always higher than that of Dutch subjects. The mean difference between the two groups seems fairly large. Comparing the prediction results (p-pred) calculated from the identification tasks 1, 3 and the obtained discrimination data from task 4 (p-obt), it seems Chinese subjects are at least correlated in a downward tendency along the continuum while this is not the case for the Dutch group. (FIGs 7~10) Besides that, the accuracy probability value is much closer in general for the Chinese group than the Dutch group in almost every step. Hence it is likely that Chinese subjects' performances in identification tasks agree with their behavior in discrimination tasks more than the Dutch peers.

---

[16] P(obt) is averaged value across 6 pairs: pair 1-3; 2-4; 3-5; 4-6; 5-7; 6-8.



FIGs 7-10 Task-wise comparison of predicted and obtained discrimination
Horizontal: 1-3; 2-4; 3-5; 4-6; 5-7; 6-8 pairs Vertical: mean accuracy probability
Left: Chinese group Right: Dutch group

In brief, the results of the CP index analysis confirm the hypothesis that Chinese listeners are more likely to perceive the lexical tonal stimuli categorically than Dutch listeners.

4 General discussion

4.1 Overall performance

The results reported in this study support the view that lexical tones in Mandarin Chinese are, to some extent, categorically perceived by native listeners, while they seem to be perceived non-linguistically by Dutch listeners. This general finding is consistent with many of the previous investigations on CP of lexical tones. Although only T1 and T4 have been tested in this experiment, it is likely that all 4 tones in Mandarin Chinese follow the same pattern[17].

The results of the identification tasks (e.g. slopes and intercepts) illustrate quite a high degree of categorization in tones among Chinese listeners, while Dutch listeners seem to lack such categories. The results of discrimination tasks are distinct from previous reports in cross-linguistic studies on lexical tones in that, on average, Dutch listeners seem to perform better than their Chinese peers in response accuracy and speed, though with high variation. Their high sensitivity to prosody, perhaps stemming from L1 stress and intonation patterns or non-linguistic tonal experiences, seems to be the reason for that. Looking from the angle of the Chinese group, it could be that the Chinese listeners'

---

[17] However, it is likely that some contrasts are more difficult to show CP effect. See Wu and Lin (2008) for a T2-T3 study.

perceptions are impeded by their native phonological category when facing small acoustic differences within the same category. RT in discrimination tasks do reveal a small advantage for the Dutch group, which might indicate that less processing time is required if the sound does not involve linguistic categorization. More accurate ways to calculate response time (e.g. eye-tracking or button box) are needed in further studies for a better interpretation of this issue. Besides that, studies may also focus on the question of how different L1 might have an influence on tone perception[18].

Finding CP solely on the basis of identification or discrimination data may not be adequate: one could even conclude that no CP is found in native tonal language speakers (Abramson, 1979; Francis et al. 2003). CP index analysis provides a meaningful interpretation of the tasks. Despite the fact that no absolute category could be easily defined, the CP index results obtained indicate some agreement between the observed probabilities from the discrimination data and the predicted probabilities from the identification data in the Chinese group. This can be seen as evidence favoring a stronger degree of CP. On the other hand, even though the intercept value from PROBIT analysis shows that the category boundary between T1 and T4 lies probably somewhere between 2.8 and 3 along the continuum, it is still not warranted to make the claim that there is a definite point (intercept / boundary location) indicating where the category boundary is. The individual data support the opinion that the CP of tones is gradient (Schouten and Van Hessen, 1992) and varies across individuals.

Psychophysical perception by Dutch listeners can be observed through their sensitivity towards the distinction between steps 1 and 2 along the continuum. The memory based account (Xu et al., 2006) can not be verified, since nonspeech stimuli are not tested in this experiment. However, the question why the Dutch group reveals such high individual variability still needs to be answered under such a cognitive approach.

4.2 Questions and further research

Some Dutch individuals perform as well as or better than their Chinese peers in identification tasks. They seem to be able to identify moderately well each stimulus along the continuum. To answer the question of how these Dutch subjects perceive lexical tones, several explanations can be addressed. One possible, yet highly unlikely, explanation is that these Dutch subjects were exposed to some tonal languages at a young age and that their perception of tones has not reorganized since. An alternative explanation could be that, like Chinese listeners, these Dutch listeners are using the prosodic information from L1 such as tone, intonation and stress, etc. for L2 tonal perception. (For example, it might be the case that Dutch listeners interpret the falling stimuli as an imperative intonation with a protocol value between stimulus 4 and 5, based on the identification results.) Indeed, Dutch is a stress language with abundant intonation and stress. However, this can hardly explain why other Dutch listeners don’t go the same way and perform poorly in the identification tasks. Besides that, none of these Dutch subjects speak a native tonal dialect of Dutch and therefore the only prosodic information they can use from L1 is intonation or stress, hardly helpful in the tasks. Although this alternative explanation seems not plausible here, it raises the intriguing question of whether Dutch listeners who speak a certain dialect with lexical tones, such as Limburgish in the

---

[18] Dutch listeners seem to perform better than English and French listeners in the experiments shown in literature review section, though making comparison under different experimental conditions across papers may not be wise.

southern part of the Netherlands, will use L1 tonal information and perform differently from and possibly better than Dutch listeners of a dialect without tones. The third explanation arises when many of these good Dutch performers self-report previous experience in non-linguistic (musical[19]) tonal variations such as playing guitar, piano or singing in a church choir. It is thus reasonable to believe that a listener's acoustic experience in non-linguistic prosody also influences his or her supra-segmental perception, probably especially when no rescue is available from their L1 inventory. Follow up research should also address this issue and set up a "musician" group in the study. By all means, the high performance of these Dutch listeners again shows the fact that the stimuli are not linguistically categorically perceived by them; instead, they use their psychophysical ability (natural non-linguistic category) to identify and discriminate lexical tones.

Another potentially interesting point lies in the fact that isolated single syllable stimuli are used in all four tasks. This should be considered adequate, since Chinese words are monosyllable based. Yet this does not tell how lexical tones are perceived in context. It is very likely that perception performance in context is different from in isolation. For example, RTs are predicted to be faster, since one study shows that short portions of tone contours are sufficient for native listeners' identification (Whalen and Xu, 1992). In the current study, however, due to the task requirements, subjects paid more attention to tonal variation and spent more time in the discrimination task. This prediction can be testified in the follow up tests[20].

5 Conclusion

In summary, a cross-linguistic approach shows that Chinese and Dutch listeners perceive tones in a different fashion. Chinese listeners perceive lexical tones more categorically than their Dutch peers: Chinese subjects' performance is biased by their L1 tonal categories, while Dutch subjects' performance is psychophysically based. Without such a categorical bias, Dutch subjects actually outperform Chinese subjects in discriminating the tonal differences, even though they do not identify tones in the same way as Chinese subjects do. This (non)L1 influence and subjects' (non-linguistic) natural auditory sensitivity in lexical tone perception, also shown in the recent literature (Francis et al., 2003; Francis et al., 2008; Kaan et al., 2008), is considered the most appropriate explanation to account for the different perception patterns and the variability between the two language groups. Chinese listeners perceive lexical tones as within their phonological categories, in the same way as they perceive phonemes, while Dutch listeners perceive them as non-linguistic tonal information.

Acknowledgements

From the beginning, this study received support from many people. I would like to thank Xiaoli Dong for providing her voice for the acoustic stimuli and Hugo Quene for the detailed instruction of PRAAT. I would like to thank all the subjects who participated in this study. I would like to thank Desiree Capel, Elise de Bree, Annemarie Kerkhoff, Maartje De Klerk for lab assistance and Theo Veenker for script generation. I would like to thank Tom Lentz, Jurriaan Witteman and Hugo Quene

---

[19] A good musical instrument player or a singer is supposed to have professional training and has higher sensitivity in distinguishing musical tonal contrasts with slight tonal differences compared to normal peers.

[20] In the cross-linguistic follow up study (identification and discrimination of lexical tones in context), tasks need to be set in such a way that they are also fair to the non-native subjects of a tonal language to perceive.

for statistics aid, and Pierre Halle for providing the PROBIT analysis tool. Finally I would like to thank Bert Schouten for very detailed CP instruction and reviewing of this paper, thank Glen Chesnut for the grammar check, and Rene Kager for supervising this paper as well as the internship.

References

Abramson, A. (1979). *The noncategorical perception of tone categories in Thai.* In B. Lindblom, & S. Ohman (Eds.), Frontiers of speech communication (pp. 127 – 134). London: Academic Press.

Best, C. T., & Strange, W. (1992). *Effects of phonological and phonetic factors on cross-language perception of approximants.* Journal of Phonetics, 20, 305 – 330.

Coster, D. C., & Kratochvil, P. (1984). *Tone and stress discrimination in normal Peking dialect speech.* In B. Hong (Ed.), New papers in Chinese linguistics (pp. 119 – 132). Canberra: Australian National University Press.

Cutting, J. E. & Rosner, B. S. (1976). *Discrimination functions predicted from categories of speech and music.* Perception & Psychophysics, 20, 87-88.

Francis, A.L., Ciocca, V., & Ng, B.K.C. (2003). *On the (non)categorical perception of lexical tones.* Perception & Psychophysics, vol. 65, no. 7, pp. 1029-1044.

Francis, A.L., Ciocca, V., Ma, L. & Fenn, K. (2008) *Perceptual learning of Cantonese lexical tones by tone and non-tonal language speakers.* Journal of Phonetics 36, 268 – 294

Gandour, J. (1978). *The perception of tone.* In Fromkin, V. A. (ed.), Tone: A Linguistic Survey, pp. 41-76. New York: Academic Press.

Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Van Lancker, D., & Hutchins, D. (2000). *A cross-linguistic PET study of tone perception.* Journal of Cognitive Neuroscience, 12, 207-222.

Gandour, J., Wong, D., & Hutchins, D. (1998). *Pitch processing in the human brain is influenced by language experience.* Neuroreport, 9, 2115-2119.

Gandour, J., Dzemidzic, M., Wong, D., Lowe, M., Tong, Y., Hsieh, L. & Lurito, J. (2001) *Differential neural circuitry underlies the perception of Chinese tone and intonation: an fMRI study.* NeuroImage 13: S533

Halle P., Chang, Y-C. & Best, C. (2004) *Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners.* Journal of Phonetics, 32, 395-421.

Henderson, A.I. & Nelms, S. (1980). *Relative Salience of Intonation Fall and Pause as Cues to the Perceptual Segmentation of Speech in an Unfamiliar Language.* Journal of Psycholinguistic Research, Vol. 9, No. 2.

Hsieh, L., Gandour, J., Wong, D. & Hutchins, G.D. (2000) *A PET study of the perception of Chinese tones.* NeuroImage 11: S313

Kaan, E., Barkley, C.M., Bao M-Z. & Wayland, R. (2008). *Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese: An event-related potentials training study.* Neuroscience, 9:53

Klein, D., Zatorre, R., Milner, B., & Zhao, V. (2001). *A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers.* Neuroimage, 13, 646-653.

Kong, Q-M. (1987) *Influence of tones upon vowel duration in Cantonese.* Language and Speech, 30, 387-399.

Lee, C-Y. (2007). *Does horse activate mother? Processing lexical tone in form priming.* Language and Speech, 50(1), 101-123

Lee, L., & Nusbaum, H. (1993). *Processing interactions between segmental and suprasegmental* *information in native speakers of English and Mandarin Chinese.* Perception and Psychophysics, 53, 157 – 165.

Macmillan, N.A., Kaplan, H.L. & Creelman, C.D. (1977). *The Psychophysics of categorical perception.* Psychological Review, 84, 452-471.

Pollack, I. & Pisoni, D. (1971). *On the comparison between identification and discrimination tests in speech perception.* Psychonomic Science, 24, 299-300.

Repp, B.H., and Lin, H.B. (1990). *Integration of segmental and tonal information in speech perception: A cross-linguistic study.* J. Phonetics 18, 481–495.

Brown-Schmidt, S. & Canseco-Gonzalez, E. (2004). *Who do you love, your mother or your horse? An event-related brain potential analysis of tone processing in Mandarin Chinese.* J Psycholinguist Res 33(2): 103-135

Schouten, M.E.H. & Van Hessen, A.J. (1992). *Modeling phoneme perception. I: Categorical perception.* Journal of the Acoustical Society of America, 92, 1841-1855.

Schwanhaeusser, B. (2008). *Lexical Tone Perception and Production: The role of language and musical background.* Dr dissertation. University of Western Sydney, MARCS Auditory Laboratories

Soto-Faraco, S., Sebastian-Galles, N., & Cutler, A. (2001). *Segmental and suprasegmental mismatch in lexical access.* Journal of Memory and Language, 45, 412 – 432.

Stagray, J.R., and Downs, D. (1993). *Differential sensitivity for frequency among speakers of a tone and a nontonal language*. J. Chin. Linguist. 21, 144–162.

Van Hessen, A.J. & Schouten, M.E.H. (1999) *Categorical perception as a function of stimulus quality.* Phonetica, 56, 56-72

Van Lancker, D., & Fromkin, V. (1973). *Hemispheric specialization for pitch and "tone": Evidence from Thai*. Journal of Phonetics, 1, 101-109.

Veenker, T.J.G. (1998). *FEP, A tool for designing and running computerized experiments.* Utrecht Institute of Linguistics – OTS, Utrecht University, Utrecht, The Netherlands.

Wang, W. S-Y. (1973). *The Chinese language*. Scientific American, 228, 50–60.

Wang, W.S.-Y. (1976). *Language Change*. In S.R. Harnad, H.D.Stek-lis, & J. Lancaster (eds.), Origins and evaluation of language and speech, Annals of the New York Academy of Sciences, vol. 280, pp. 61-72.

Wang, Y., Jongman, A., & Sereno, J. A. (2001). *Dichotic perception of Mandarin tones by Chinese and American listeners.* Brain and Language, 78, 332-348.

Wang, Y., Behne, D. M., Jongman, A., & Sereno, J. A. (2004). *The Role of Linguistic Experience in the Hemispheric Processing of Lexical Tone.* Applied Psycholinguistics, vol. 25, no. 3, pp. 449-466.

Whalen, D., & Xu, Y. (1992). *Information for Mandarin tones in the amplitude contour and in brief segments.* Phonetica, 49, 25 – 47.

Wu, X.H. & Lin, H. (2008) *Perception of Mandarin Tones by Mandarin and English Listeners*. Journal of Chinese Language and Computing, 18(4), 175-187

Xu, Y., Gandour, J.T. & Francis, A.L. (2006) *Effects of language experience and stimulus complexity on the categorical perception of pitch direction.* Acoustical Society of America, Vol.120, No.2, 1063-1074

Appendix II – F0 value in 4 points from /ta1/ to /ta4/

In these tables, /ta1/ = stimulus 1 while /ta4/ = stimulus 8. The rest numbers correspond to the stimuli created along the continua. The four letters (A, B, C and D) represent four continua created from different original /ta1/ - /ta4/ sounds produced by the same speaker.

	Starting Point	Interpolate 1	Interpolate 2	Ending Point
A1	288.300	289.000	289.000	290.000
A2	295.614	291.857	275.700	271.484
A3	302.928	294.714	262.425	252.970
A4	310.242	297.571	249.140	234.456
A5	317.556	300.428	235.855	215.942
A6	324.870	303.285	222.570	197.428
A7	332.184	306.142	209.285	178.914
A8	339.500	309.000	196.000	160.400

Continuum A

	Starting Point	Interpolate 1	Interpolate 2	Ending Point
B1	275.500	277.633	279.766	281.900
B2	283.582	282.443	267.202	264.592
B3	291.664	287.253	255.635	247.281
B4	299.746	292.063	244.068	229.970
B5	307.828	296.873	232.501	212.659
B6	315.910	301.683	221.934	195.348
B7	323.992	306.493	210.367	178.037
B8	332.077	311.300	198.800	160.726

Continuum B

	Starting Point	Interpolate 1	Interpolate 2	Ending Point
C1	278.795	276.670	274.525	272.390
C2	293.044	284.788	259.178	255.264
C3	307.293	292.906	243.829	238.136
C4	321.542	301.024	228.480	221.008
C5	335.791	309.142	213.131	203.880
C6	350.040	317.260	197.782	186.752
C7	364.289	325.378	182.433	169.624
C8	378.537	333.494	167.084	152.496

Continuum C

	Starting Point	Interpolate 1	Interpolate 2	Ending Point
D1	272.502	272.826	273.150	273.474
D2	287.777	278.903	259.092	256.802
D3	303.052	284.980	245.032	240.133
D4	318.327	291.057	230.972	223.464
D5	333.602	297.134	216.912	206.795
D6	348.877	303.211	202.852	190.126
D7	364.152	309.288	188.792	173.457
D8	379.430	315.366	174.732	156.788

Continuum D