



HAL
open science

The development of tone discrimination in infancy: Evidence from a cross-linguistic, multi-lab report

Marina Kalashnikova, Leher Singh, Angeline Tsui, Eylem Altuntas, Denis Burnham, Ryan Cannistraci, Ng Bee Chin, Ye Feng, Laura Fernández-Merino, Antonia Götz, et al.

► To cite this version:

Marina Kalashnikova, Leher Singh, Angeline Tsui, Eylem Altuntas, Denis Burnham, et al.. The development of tone discrimination in infancy: Evidence from a cross-linguistic, multi-lab report. *Developmental Science*, In press, 10.1111/desc.13459 . hal-04322749

HAL Id: hal-04322749

<https://hal.science/hal-04322749>

Submitted on 5 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**The development of tone discrimination in infancy: Evidence from a cross-linguistic,
multi-lab report**

Kalashnikova, M.^{1,2*†}, Singh, L.^{3*}, Tsui⁴, A., Altunas⁵, E., Burnham, D.⁵, Cannistraci, R.^{6,7}, Chin, N. B.⁸, Feng, Y.^{9,10,11}, Fernández-Merino, L.^{1,12}, Götz, A.^{5,13}, Gustavsson, L.¹⁴, Hay, J.⁶, Höhle, B.¹³, Kager, R.^{15,16}, Lai, R.^{9,10}, Liu, L.^{5,17,18,19}, Marklund, E.¹⁴, Nazzi, T.^{20,21}, Oliveira, D. S.⁶, Olstad, A. M. H.¹⁹, Picaud, A.^{20,21}, Schwarz, I.-C.^{14,22}, Tsao, F.^{23,24}, Wong, P. C. M.^{9,10}, Woo, P. J.²⁵

1. Basque Center on Cognition, Brain and Language, San Sebastian, Spain
2. Ikerbasque, Basque Foundation for Science, Bilbao, Spain
3. Department of Psychology, National University of Singapore, Singapore
4. Stanford University, Stanford, CA, USA
5. MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Australia
6. Department of Psychology, University of Tennessee, Knoxville, USA
7. Department of Psychology, Lafayette College, USA
8. Linguistics and Multilingual Studies, School of Humanities, Nanyang Technological University, Singapore
9. Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, Hong Kong SAR, China
10. Brain and Mind Institute, The Chinese University of Hong Kong, Hong Kong SAR, China
11. Department of Linguistics, Beijing Language and Culture University, Beijing, China
12. University of Basque Country, San Sebastian, Spain
13. Linguistics Department, University of Potsdam, Germany
14. Department of Linguistics, Stockholm University, Stockholm, Sweden
15. Department of Languages, Literature and Communication, Utrecht University, The Netherlands
16. Utrecht Institute of Linguistics OTS, Utrecht University, The Netherlands
17. School of Psychology, Western Sydney University, Penrith, Australia
18. Centre of Excellence for the Dynamics of Language, Australian Research Council. Canberra, Australia
19. Center of Multilingualism across the Lifespan, University of Oslo, Oslo, Norway
20. INCC, CNRS Paris, France
21. Université Paris Cité, Paris, France
22. Department of Special Education, Stockholm University, Stockholm, Sweden
23. Department of Psychology, National Taiwan University, Taiwan
24. Imaging Center for Integrated Body, Mind and Culture Research, National Taiwan University, Taiwan
25. Department of Psychology, Sunway University, Malaysia

*Kalashnikova and Singh share the first author contribution to this work.

†Corresponding author: Dr. Marina Kalashnikova, Basque Center on Cognition, Brain and Language, Paseo Mikeletegi 69, San Sebastian, 20018, Spain

Data availability statement: All materials and data can be accessed at

https://osf.io/t6sp2/?view_only=b731a1b1c17643e98593155aeae9e626 or by request to the corresponding author Marina Kalashnikova at m.kalashnikova@bcbl.eu.

Conflict of interest statement: The authors declare no conflicts of interest related to this work.

Funding sources:

We thank the following funding sources for supporting this work. In Spain, MK received support from the Basque Government through the BERC 2022-2025 program, by the Spanish State Research Agency through BCBL Severo Ochoa excellence accreditation CEX2020-001010/AEI/10.13039/501100011033, and by the Ramon y Cajal Research Fellowship, RYC2018-024284-I. In Singapore (NUS), LS received support from the ODPRT grant for research excellence. In Australia, EA received support from the MARCS Institute for Brain, Behaviour and Development Research Funds, Western Sydney University; and DB received support from the MARCS Institute for Brain, Behaviour and Development Research Funds, Western Sydney University. In Singapore (NTU), BC received support from the Nanyang Technological University, Singapore. In Germany, BH and AG received support from the DFG (German Research Foundation) FOR 2253 HO 1960/19-2. In Sweden, LG received support from the Riksbankens Jubileumsfond P17-0175; EM received support from the Magnus Bergvall foundation 2018-02855, Stiftelsen Marcus och Amalia Wallenbergs minnesfond 2019.0030; and I.-C.S. received support from the Stockholm University Black Box SU FV-5.1.2-2875-18 to I.-C.S. In Norway, LL received support from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 798658 hosted by the Center for Multilingualism across the Lifespan at the University of Oslo, financed by Research Council of Norway through its Centers of Excellence funding scheme grant agreement No. 223265 and Western Sydney

TONE DISCRIMINATION IN INFANCY

University School of Psychology 20820 83181. In France, TN and AP received support from the LABEX EFL (ANR-10-LABX-0083). In Taiwan, FT received support from the National Science and Technology Council, Taiwan (110-2410-H-002 -128 -MY2). In Hong Kong, PW received support from the University Grants Committee (HKSAR; RGC34000118, C4055-19G) and the Innovation and Technology Fund (HKSAR; ITS/067/18). In Malaysia, PJW received support from the Sunway University Internal Grant, INT-2019-SST-PSY-01.

Acknowledgments:

We are deeply grateful to all the infants who took part in this study and their families. We also thank Elena Aguirrebengoa, Patricia Jimenez, Annabel Tan, Ella Ward, Chunzi Li, Kin Man Carmen Tang, Tom Fritzsche, Floor Braun, Desiree Capel, Charlotte Koevoets, Klara Marklund Hjerpe, Anna Ericsson, Petter Kallioinen, Bente Sand Aronsson, Glenneze Ong, and Ng J-Wen, and Anne Karin Gisvold for assistance with data collection across the participating laboratories. In addition, the Hong Kong team would like to thank The Chinese University of Hong Kong (CUHK) – Utrecht University (UU) Joint Centre for Language, Mind and Brain, the CUHK – NTU – WSU Joint Laboratory for Infant Research, and participants of the Stanley Ho Developmental Cohort Study. The Norway team thanks Elizabeth Lanza, Fredrik Eugen Christiansen, Ingeborg Sophie Ribu, Hanne Gram Simonsen for help in establishing the infant laboratory and setting up the testing environment.

Research Highlights

- This is a multi-language and multi-lab investigation of young infants' ability to discriminate lexical tones.
- This study included data from 13 laboratories testing 5-, 10-, and 17-month-old monolingual (tone, pitch-accent, non-tone) and bilingual (tone/non-tone, non-tone/non-tone) infants.
- Overall, infants discriminated a perceptually similar and a distinct non-native tone contrast, although there was no evidence of a native tone-language advantage in discrimination.
- These results demonstrate maintenance of tone discrimination throughout development.

Abstract

We report the findings of a multi-language and multi-lab investigation of young infants' ability to discriminate lexical tones as a function of their native language, age, and language experience, as well as of tone properties. Given the high prevalence of lexical tones across human languages, understanding lexical tone acquisition is fundamental for comprehensive theories of language learning. While there are some similarities between the developmental course of lexical tone perception and that of vowels and consonants, findings for lexical tones tend to vary greatly across different laboratories. To reconcile these differences and to assess the developmental trajectory of native and non-native perception of tone contrasts, this study employed a single experimental paradigm with the same two pairs of Cantonese tone contrasts (perceptually similar vs. distinct) across 13 laboratories in Asia-Pacific, Europe, and North-America testing 5-, 10-, and 17-month-old monolingual (tone, pitch-accent, non-tone) and bilingual (tone/non-tone, non-tone/non-tone) infants. Across the age range and language backgrounds, infants who were not exposed to Cantonese showed robust discrimination of the two non-native lexical tone contrasts. Contrary to this overall finding, the statistical model assessing native discrimination by Cantonese-learning infants failed to yield significant effects. These findings indicate that lexical tone sensitivity is maintained from 5 to 17 months in infants acquiring tone and non-tone languages, challenging the generalisability of the existing theoretical accounts of perceptual narrowing in the first months of life.

Keywords: Lexical tone; Speech discrimination; Perceptual reorganization; Bilingualism; Infancy

TONE DISCRIMINATION IN INFANCY

During their first postnatal year, infants face the challenging task of learning the phonological categories of their native language(s), an ability that is integral to early language development. To date, evidence for the development of native phoneme categories is based mainly on monolingual infants acquiring Germanic or Romance languages. Most of these languages *primarily* use consonants and vowels to differentiate lexical meaning. Such consonant-vowel-only languages are in the minority; over half of the world's languages are tone or pitch-accent languages that use, in addition to consonants and vowels, pitch variations within words to differentiate word meaning (Maddieson, 2008; Yip, 2002). These languages are widely spoken around the world and are native to the majority of language learners (Fromkin, 1978). In spite of this, these languages have not been widely studied in the language acquisition literature.

The inclusion of tonal (lexical tone and pitch-accent) languages in language acquisition research promises to enrich our understanding of early language development in several ways. First, an investigation of tone acquisition affords an opportunity to understand an aspect of language learning that applies to the majority of children around the world who learn a tone language as their native language. Second, it allows us to test whether prevailing models of infant speech perception generalise to tone languages. Third, tone is linguistically interesting as it is carried by a dimension of speech – vocal pitch – that varies in communicatively significant ways in all languages, yet it is only lexicalised in tone languages. The differentiation of pitch into its various functions is a challenge for every learner but is arguably a particularly interesting challenge for tone language learners who must isolate lexical tone from other sources of pitch variation such as intonation (see Kager, 2018). Finally, there is emerging evidence that tone language learners differ from non-tone language learners not only in their perception of lexical pitch but also in their sensitivity to vowels and consonants (Gómez, Mok, Ordin, Mehler, & Nespors, 2018; Wiener & Turnbull,

TONE DISCRIMINATION IN INFANCY

2016; Poltrock, Chen, Kwok, Cheung, & Nazzi, 2018; Wewalaarachchi, Wong & Singh, 2017; Chen, Lee, Luo, Lai, Cheung, & Nazzi, 2021). Therefore, studying tone perception in tone and non-tone language learners has the potential to reveal different facts about how more-commonly studied properties of language – such as vowels and consonants – may be constrained by each learner’s language background. These arguments have inspired a recent surge in research on lexical tone acquisition in infancy and early childhood (e.g., see Singh & Fu, 2016 for an overview). In particular, there has been recent interest in investigating how infants discriminate lexical tones. This research has yielded valuable but somewhat contradictory findings. The aim of this study is to provide the first comprehensive cross-linguistic investigation of infants’ early ability to discriminate lexical tones and to systematically measure how this ability is impacted by age, language background, language experience, and the psychoacoustic properties of specific lexical tone contrasts.

Infants come to the world equipped with the ability to discriminate many phonetic contrasts across the world’s languages. Although newborns have certain acoustic/phonetic sensitivities (Burnham, 1986; Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Polka & Werker, 1994; Werker & Tees, 1984a), these are not language-specific, so infants are born prepared to learn any language (Kuhl, 2004). During their first year, infants undergo a period of perceptual reorganisation or perceptual narrowing, which is characterised by maintenance or an increase in sensitivity to native speech contrasts and a corresponding decrease in sensitivity to many non-native speech contrasts. This is proposed to occur around the age of four to six months for vowel categories and around nine to 11 months for consonants (Best, 1994; Bosch & Sebastián-Gallés, 2003; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Kuhl et al., 2006; Polka & Bohn, 2011; Polka & Werker, 1994; Stager & Werker, 1997; Werker & Tees, 1984a). These age ranges are not critical periods after which native speech perception abilities are irreversibly consolidated (Werker, 2018; Werker & Hensch,

2015). Rather, this reorganisation may involve a change in processing biases but not loss of sensorineural responsivity, as adults can still discriminate non-native speech contrasts when short inter-stimulus durations are used or after training (Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Götz, Yeung, Krasotkina, Schwarzer, & Höhle, 2018; Mazuka, Hasegawa, & Tsuji, 2014; Werker & Logan, 1985; Werker & Tees, 1984b). Even in infancy, sensitivity to some native but phonetically non-salient consonant contrasts continues to develop during the first year (Narayan, Werker, & Beddor, 2010; Tsao, Liu, & Kuhl, 2006), and renewed sensitivity to a non-native contrast can be observed after exposure to the non-native language in which the contrast is used (Kuhl, Tsao, & Liu, 2003).

The Phonology of Tone Languages

The phonemic inventories of tone languages include consonants, vowels, and lexical tones. Acoustic analyses of lexical tones show a number of distinguishing acoustic properties (duration, amplitude, vocal range, and voice quality). However, they are primarily defined by modulations in pitch (fundamental frequency, F₀) height and contour (Liu & Samuel, 2004). As tones span an entire syllable in most tone languages, they are often classified as suprasegmental cues, despite their function as phonological segments that distinguish between words. Similar to tone languages, pitch-accent languages (e.g., Japanese, Norwegian, and Swedish) also employ systematic changes in pitch at the lexical level, but in a more restricted manner than in lexical tone languages (see Hyman, 2009 for a discussion). Specifically, in these languages, pitch changes refer to the relative pitch between adjacent syllables, and they are only imposed on a subset of syllables rather than on every syllable of the word, and pitch changes are not obligatorily assigned to all words or all syllables of the language.

As is the case with consonants and vowels, the composition of tone and to a lesser extent pitch-accent inventories vary widely across tone languages. In this study, Cantonese

TONE DISCRIMINATION IN INFANCY

tones were used as stimuli (Figure 1 presents the Cantonese tone inventory), and infants acquiring tone languages (e.g., Cantonese, Mandarin Chinese), pitch-accent languages (Norwegian, Swedish), and non-tonal languages (Basque, Dutch, English, French, German, and Spanish) participated in this study. Lexical tones are conventionally described using Chao numerals (Chao, 1968), which refer to a method of coding the pitch onset, offset and inflection points of a tone according to a scale that ranges from 1 (lowest pitch) to 5 (highest pitch).

Perceptual Reorganisation for Tones

While research on infant discrimination of native and non-native consonant and vowel categories has mainly provided evidence for facilitation or maintenance for native contrasts and attenuation for non-native contrasts (see Aslin & Pisoni, 1980 for a theoretical account of these developmental pathways), evidence from studies of infants' lexical tone discrimination has been mixed. Three distinct developmental patterns have been found both for native and non-native lexical tone discrimination: *facilitation or maintenance* over age, *attenuation* over age, and non-monotonic discrimination profiles over age, i.e., *resurgence*. Moreover, different patterns have been found for monolingual and bilingual infants and for tone contrasts that differ in psychoacoustic salience (see Singh & Fu, 2016 for a review). Figure 2 provides a summary of studies of the developmental patterns of tone discrimination in infants in terms of evidence for *facilitation*, *maintenance*, *attenuation*, and *resurgence* for monolingual and for bilingual infants, along with the Chao values of the tone contrasts used in each study and classification of the psychoacoustic salience of the contrasts as 'distinct' for high salience and 'subtle' for low salience. Below, we present a corresponding review of infant tone discrimination studies in three sections: (i) results for infants being raised in a monolingual tone or non-tone language environment, (ii) results for infants being raised in a

TONE DISCRIMINATION IN INFANCY

bilingual language environment, and (iii) the effects of psychoacoustic salience on the developmental patterns of tone discrimination.

Tone Discrimination in Monolingual Infants

Maintenance or facilitation in tone language infants. Maintenance of tone discrimination over the first year has been found in native tone learners: there is stability in Mandarin/Cantonese infants' discrimination of Thai tones between 6 and 9 months (Mattock & Burnham, 2006), in Mandarin infants' discrimination of a Mandarin tone contrast between 6-8 and 10-12 months (Tsao, 2017), and in Mandarin and Cantonese-exposed infants' discrimination of Cantonese tones between 4 and 9 months (Yeung, Chen, & Werker, 2013). Additionally, Yeung et al. (2013) reported evidence for differences in Cantonese tone perception by Cantonese and Mandarin infants even at 4 months suggesting some early influence of infants' native language experience on their tone discrimination abilities. Thus far, studies demonstrating stability have used Thai or Cantonese tones.

Using Mandarin tones, facilitation has been found in three studies on Mandarin-monolingual infants' discrimination of native Mandarin tones: Singh et al. (2018) in combination with Singh, Poh, and Fu (2016) found improvement between 6 and 12 months. Facilitation in native Mandarin tone discrimination with age was also reported by Tsao (2017) between 7 and 11 months.

Facilitation in non-tone language infants. Facilitation of tone discrimination over age has also been reported in non-tone language learning infants, predominantly using Mandarin tones. Chen and colleagues found an improvement over age in Dutch-learning infants' tone discrimination: infants did not discriminate a subtle Mandarin contrast at 4 months, but they did do so at 6 and 12 months of age (Chen & Kager, 2016; Chen, Stevens & Kager, 2017). Similarly, testing English language infants on Mandarin tones, Singh et al. (2018) found that infants progressed from discriminating neither distinct nor subtle Mandarin

TONE DISCRIMINATION IN INFANCY

tone contrasts at 6 months to discriminating both types of tone contrasts at 12 months. Tsao (2017) also demonstrated an improvement in the perception of Mandarin tones between 6-8 and 10-12 months in English-learning infants.

Attenuation in non-tone language infants. No studies have reported evidence of attenuation to tone in tone language learning infants. In contrast, evidence for attenuation of tone discrimination in non-native tone perception has been found in five studies using Thai, Mandarin, and Cantonese tones. Mattock and Burnham (2006) compared discrimination of a Thai tone contrast by non-tone (English) and tone (Mandarin or Cantonese) learning infants at 6 and 9 months. They reported maintenance for both contrasts by the tone language infants and attenuation by the non-tone language infants. Similar results of maintenance in tone (Mandarin) infants and attenuation in non-tone (French) infants were found for Thai tones (Mattock, Molnar, Polka, & Burnham, 2008; Cabrera et al., 2015). Similar attenuation was found for a Mandarin tone contrast in French-learning infants between 4 and 11 months (Shi, Santos, Gao, & Li, 2017). Using Cantonese tone stimuli, Yeung et al. (2013) found similar attenuation between 4- and 9-month-old non-tone English monolingual infants.

Maintenance or resurgence in tone and non-tone language infants. One investigation of non-tone language learners showed maintenance for a Mandarin contrast by French monolingual infants across 4, 8, and 11 months (Shi et al., 2017), and another showed maintenance of a Mandarin contrast by Dutch infants from 5 to 17 months (Liu & Kager, 2014). In addition, maintenance has been found for tone language infants listening to a non-native tone language: Cabrera et al. (2015) found maintenance of discrimination of a non-native Thai tone contrast in 6- and 10-month-old Mandarin infants, and Ramachers, Brouwer, and Fikkert (2018) found maintenance of discrimination by Dutch 6- to 12-month-old infants tested on a contrast from Limburgian, which is a restricted tone (i.e., pitch-accent) language.

tone discrimination in infancy

Several studies raise the possibility that developmental modulation of tone discrimination in infants' first year of life may be transient, and that further perceptual shifts may be observed in their second year. As described above, Liu and Kager (2014) found that Dutch-learning infants maintain discrimination of a Mandarin tone contrast between 5 and 18 months of age. However, when tested on a more subtle contrast resulting from compressing the acoustic distance between the tone pair, a U-shaped pattern emerged with discrimination at 5-6 months, then a decline in discrimination between 8 and 12 months, and then a resurgence of discrimination at 14-15 months with a further increase at 17-18 months. A similar pattern has recently been found for German infants' discrimination of a Cantonese tone contrast (the same contrast as used by Yeung et al., 2013). In this study, there was an attenuation in discrimination between 6 and 9 months but then a resurgence of discrimination at 18 months (Götz et al., 2018; see also Hay, Cannistraci, & Zhao, 2019; Hay, Graf-Estes, Wang, & Saffran, 2015; and Singh, Tam, Chan, & Golinkoff, 2014 for evidence of Mandarin tone discrimination in word-learning tasks by English-learning infants at 14, 18, and 19 months respectively and Liu & Kager, 2018 for evidence of similar tone sensitivity in 14-month-old Dutch infants).

Tone Discrimination in Bilingual Infants

In addition to monolingual learners, recent research has found an impact of bilingual experience on tone perception. Liu and Kager (2017) tested bilingual infants who were acquiring Dutch and another non-tone language from 5 to 18 months of age using a Mandarin tone contrast. Similar to their findings with monolingual infants (see above, Liu & Kager, 2014), when tested on an artificially compressed contrast, bilingual infants also demonstrated a U-shape developmental pattern, but the resurgence in discrimination was observed earlier for bilingual (11-12 months) than for monolingual infants (17-18 months, Liu & Kager, 2014). The authors suggested that bilinguals' complex linguistic environment may facilitate

TONE DISCRIMINATION IN INFANCY

greater acoustic sensitivity, even for lexical tones, although these are not lexically contrastive in the bilinguals' native languages (see also Liu & Kager, 2016).

Two studies have investigated tone discrimination in bilingual samples learning one tone language and one non-tone language. Singh and colleagues (2018) investigated the development of tone perception in tone/non-tone (Mandarin-English) bilinguals and found that they did not discriminate either subtle or distinct tone contrasts at 6, 9, or 12 months. This difficulty may be due to the conflicting lexical role pitch plays in English (pitch is used to draw non-lexical intonational distinctions) and in Mandarin (pitch is used to draw non-lexical *and* lexical distinctions). It is important to note that their task did not include any language-identifying cues, which invites the possibility that tone/non-tone bilingual infants may require contextual support to differentiate pitch variation in a language-appropriate manner. Support for this possibility comes from a study with English-Mandarin bilinguals that tested sensitivity to tone in a word segmentation task that carried language-identifying cues, in which infants demonstrated language-selective pitch differentiation by 11 months (Singh & Foong, 2012). One other study assessed tone discrimination in tone/non-tone bilinguals using Limburgian tones and found maintenance of discrimination between 6 and 12 months (Ramachers et al., 2018). However, the sample size in this hard-to-recruit population was small and as such may not have had sufficient power to detect age-dependent changes.

Psychoacoustic Salience of Tone Contrasts

The above developmental patterns for tone discrimination are qualified by the acoustic properties of the tone contrasts. As can be seen in Figure 2, previous studies vary widely in the tone languages chosen for stimuli, and within these languages, studies vary in their selection of particular tone contrasts. One factor that has been manipulated in many studies is *psychoacoustic salience* of tone pairs. There is little agreement regarding the

TONE DISCRIMINATION IN INFANCY

specific properties that would qualify a tone contrast as distinct or subtle and thus perceptually difficult or easy. For instance, it has been proposed that non-native tones that resemble commonly occurring prosodic patterns in infants' native language (e.g., the rising tone is similar to the rising intonation in questions) can facilitate discrimination (Hay et al., 2019). Other studies have categorised tones according to their acoustic trajectory suggesting that contour tones (height and pitch trajectory cues) are easier to discriminate than level tones (height only) (Mattock & Burnham, 2006). Still, other studies have used adults' discrimination of individual tone pairs as a basis for determining the psychoacoustic salience of tone contrasts (Singh et al., 2018). All of these are considered below.

Distinct (salient) tone contrasts. Studies that have tested discrimination of distinct tone contrasts have reported maintenance and facilitation patterns in non-tone language-learning infants with only one exception. Two studies have shown facilitation across the first year of life in English and French infants (Shi et al., 2017; Tsao, 2017), and one study reported maintenance for Dutch monolingual infants (Liu & Kager, 2014) and another for bilingual infants (Liu & Kager, 2017). In the case of native discrimination in tone-learning infants, Shi et al. (2017) reported maintenance in Mandarin infants from 4 to 14 months and two studies reported facilitation across age in the same language group (Singh et al., 2018; Tsao, 2017). However, Mattock and Burnham (2006) reported attenuation of discrimination of a distinct Thai contrast by non-tone (English) learning infants from 6 to 9 months.

Subtle contrasts. Regarding discrimination of subtle contrasts, evidence can be found in support of each of the four developmental patterns: maintenance, facilitation, attenuation, and resurgence. First, in the case of non-native discrimination, three studies have reported continuous facilitation across age in Dutch- and English-learning infants (Chen & Kager, 2015; Chen et al., 2017; Singh et al., 2018), and three studies reported resurgence for monolingual infants learning Dutch, German, as well as non-tone bilinguals (Liu & Kager,

2014; 2017; Götz et al., 2018). In contrast, five studies have reported attenuation of discrimination for subtle non-native contrasts for English- (Mattock & Burnham, 2006; Yeung et al., 2013) and French-learning infants (Cabrera et al., 2015; Mattock et al., 2008; Shi et al., 2018). In the case of native discrimination, for tone-learning infants, Singh et al. (2018) reported facilitation for the subtle rising vs. dipping contrast, and Tsao (2017) reported maintenance using the same contrast. Maintenance was also reported by Shi et al. (2018) and Tsao (2017) for Mandarin infants tested on the rising-falling contrast.

To summarise, studies have found different patterns of discrimination for distinct versus subtle contrasts (e.g., Liu & Kager, 2014; Singh et al., 2018; Tsao, 2017). There is generally greater convergence of results from studies using psychoacoustically distinct tone contrasts – these contrasts generally result in maintenance or facilitation patterns. In contrast, studies using subtle contrasts have yielded more variable patterns of discrimination. These subtle contrasts may be more vulnerable to the effects of native language background, age, and other individual factors. This provides an impetus for investigating both perceptually distinct and subtle tone contrasts in this study.

Summary and the Current Study

In sum, patterns of tone discrimination in infant speech perception are far less categorical than has been established for vowels and consonants. There is little consensus as to whether tones are subject to perceptual narrowing, which is widely considered to be a hallmark of infant language development. As can be seen in Figure 2, age-related changes in tone discrimination vary both within tone- and non-tone language learners, between facilitation, attenuation, maintenance, and resurgent development, and discrimination abilities are further modified by the psychoacoustic salience of the tone stimuli used in each study. The effects of infants' age are less predictably tethered to their tone discrimination abilities than one might expect based on previous findings with consonant and vowel discrimination.

TONE DISCRIMINATION IN INFANCY

Therefore, there is a great deal of uncertainty about whether previous accounts of perceptual reorganisation accommodate lexical tones as well as they do for consonants and vowels. This provides a strong impetus for further research across populations and languages to determine what factors constrain tone perception in infancy.

We initiated a large international research collaboration to investigate the role of experience in lexical tone perception. Our project was inspired by recent collaborative initiatives across multiple labs (e.g., Frank et al., 2017; Klein et al., 2014), and followed the example of recent multi-site infancy studies initiated by the Many Babies Consortium (e.g., Many Babies Consortium, 2020; Byers-Heinlein et al., 2021; Frank et al., 2017). This study brings together evidence on early tone perception from infant laboratories located in the Asia-Pacific region, rich in tone languages, as well as Western Europe and the United States where non-tone language monolingual and multilingual populations of infants can be accessed. Specifically, we studied infant tone discrimination across different age groups and language backgrounds, including infants acquiring non-tone, tone, and intermediate tone languages (i.e., pitch-accent languages). Therefore, in this study, infants' ability to discriminate lexical tones was compared across three language types (tone, pitch-accent, and non-tone), nine languages (Basque, Cantonese, Dutch, English, French, German, Mandarin, Norwegian, Spanish and Swedish), two language-background groups (monolingual, bilingual), three age groups (5 months, 10 months and 17 months), and two tone contrast types (perceptually similar, perceptually distinct).

In the current work, Cantonese was selected as the stimulus language. Cantonese is characterised by a complex lexical tone system with a dense lexical space comprising six tones (Figure 1). Specifically, the 25 (rising), 21 (falling), and 33 (mid-level) tones were selected for this study to assess discrimination of two types of tone contrasts: a perceptually similar contrast (25 vs. 21) and a perceptually distinct contrast (25 vs. 33). Our categorisation

TONE DISCRIMINATION IN INFANCY

of the similarity of the tone pairings was derived based on existing evidence from infant (Hay et al., 2019) and adult tone discrimination (Wang, Spence, Jongman, & Sereno, 1999). The rising vs. falling contrast was classified as perceptually similar given that these two tones overlap in their onset F0 level (So & Best, 2010). The rising vs. mid-level contrast was classified as perceptually distinct given that it comprises two tones that do not overlap in their onset level or trajectory. We also included infants from three age groups, which correspond to the ages when previous studies have identified the four developmental patterns described in our literature review. Therefore, our comprehensive design extends the study of lexical tone perception during and beyond the first year of life. The specific predictions for this study are described below.

Main Predictions

Age. Our first aim was to define the developmental trajectory for lexical tone discrimination in infants from 5 to 17 months of age. Based on previous research, we have constructed separate predictions for native and non-native discrimination patterns.

Native tone perception: Cantonese infants' discrimination of Cantonese tones.

Following the perceptual reorganisation account concerning native speech perception, Cantonese infants were proposed to show successful discrimination of the two tone contrasts across ages. An improvement over age was also predicted (Singh et al., 2016; 2018; Tsao, 2017). Additionally, following previous studies on Mandarin infants' discrimination of native tone contrasts (Singh et al., 2018; Tsao, 2008; 2017), it was possible that more successful discrimination of the perceptually distinct compared to the perceptually similar contrast would be observed for the two younger age groups.

Non-native tone perception. As summarised in this literature review and Figure 2, three performance patterns could be predicted for the development of non-native tone perception across the three age groups included in this study. The first is a decline in

tone discrimination in infancy

sensitivity to tone contrasts as a function of age (Cabrera et al., 2015; Mattock & Burnham, 2006; Mattock et al., 2008), which would replicate previous findings for most vowel and consonant contrasts (Werker & Tees, 2005). The second is maintenance or improvement in tone sensitivity, which would support the view that regardless of infants' native phonological inventory, their tone sensitivity is maintained or even improved reflecting maturation of acoustic sensitivity to pitch information in the speech signal (Singh et al. 2018; Tsao, 2017) and/or general cognitive growth (Chen & Kager, 2016). The third possible outcome is that a non-linear U-shaped pattern would be observed, which would support the proposal that infants' initial sensitivity to lexical tones decreases as a function of attunement to their native language, but later re-emerges as a product of developing general auditory processing abilities and/or the acquisition of native supra-segmental competence (i.e., native intonation patterns that can resemble the acoustic contours of lexical tones (Götz et al., 2018; Hay et al., 2019; Liu & Kager, 2014).

Language type. Our second aim was to compare tone discrimination as a function of infants' language type. Language types were defined based on the extent to which infants' native language employs pitch information to mark phonetic distinctions: tone languages, pitch-accent languages and non-tone languages. In the case of non-native tone contrasts, a decline in sensitivity was expected over age, in line with the perceptual reorganisation account, regardless of infants' native language, and the acoustic salience of specific tone contrasts (e.g., Mattock & Burnham, 2006; Mattock et al., 2008; Yeung et al., 2013). However, given more recent cross-linguistic evidence (Figure 2), we expected that language background and properties of tone pairs would mediate infants' performance in our task. First, we predicted that infants' tone sensitivity would be shaped by the role of pitch as a source of lexical contrast in infants' native languages. In this case, a decline in sensitivity during the first year of life was expected in monolingual infants from non-tonal language

backgrounds (Dutch, English, French, German), but maintained or improved sensitivity was expected in infants from tone language backgrounds (Mandarin; Shi et al., 2017). We also predicted that sensitivity would be maintained, but to a lesser extent, in infants from pitch-accent languages (Norwegian, Swedish). To our knowledge, no previous studies have assessed non-native tone perception in pitch-accent-language infants, so our prediction was based on evidence from adult research (Burnham et al., 2015).

Language background. Bilingualism has been proposed to impact the perceptual attunement trajectory for native and non-native categories, leading to the view that bilingual exposure leads to protracted periods of neural commitment to native phonetic contrasts and protracted periods in the ability to discriminate non-native phonetic contrasts (Ferjan Ramírez, Ramírez, Clarke, Taulu, & Kuhl, 2017; Petitto et al., 2012), but findings regarding the effects of bilingualism on tone perception have been mixed (Liu & Kager, 2017; Singh et al., 2018). To evaluate the effects of bilingualism on infants' performance, the monolingual and bilingual groups in this study were compared collapsing across infants acquiring tonal and non-tonal languages. Our exploratory analyses (see below) aimed to further qualify this effect by assessing the influence of tone-language experience on bilingual infants' performance in this task.

Tone contrast. Finally, the effects of tone salience on discrimination abilities across language and age groups were assessed. In light of the lack of convergence in previous findings regarding the effects of perceptual salience on infants' ability to discriminate non-native tone contrasts, we abstained from specific predictions for each language background and age group in this study. We expected that robust conclusions about stimulus effects would be drawn from this large and diverse sample. Our overall objective was to trace the developmental trajectory of each group's ability to discriminate the similar and distinct tone contrast in order to delineate the patterns in which this factor interacts with infants'

individual linguistic background, age-related maturation of acoustic sensitivity, and growing language-specific competence.

Exploratory Analyses

In addition to the main predictions listed above, our design allowed us to conduct an exploratory analysis to further assess the effects of bilingual infants' individual language experience on tone discrimination. This analysis accounted for infants' individual degree of exposure to their two languages and one or none of infants' native languages being a tone language¹. Bilingual infants acquiring a tone language might exhibit limited tone sensitivity given that our task did not provide cues to language context as in past instantiations of the task (Singh et al., 2018).

Method

Project Time Frame and General Participating Protocols

Individual labs confirmed their participation in October 2018. A power analysis (reported in Supplementary Information, SI1) indicated that a minimum sample of 177 infants was required for our planned analyses. Data collection commenced in February 2020 and was completed in April 2022². Detailed information about the standardisation of participant recruitment and data collection procedures is available in SI2. The registration protocol for this study can be accessed at osf.io/t6sp2.

Participants

¹ Our registered predictions also included bilingual infants acquiring two tone languages (Cantonese-Mandarin bilinguals). However, despite our efforts to recruit this population across our testing locations, families with these infants did not volunteer for our study. Therefore, we have removed these predictions, and instead we now discuss the inclusion of this language background group as an important direction for future research.

² The study duration exceeded the planned period of 18 months due to the COVID-19 pandemic. Individual labs interrupted data collection for varying periods of time depending on the quarantine and movement restriction measures in each country and due to interruptions in participant recruitment.

TONE DISCRIMINATION IN INFANCY

The final sample comprised a total of 462 infants (237 male, 225 female), a third tested in each of the three age groups (5, 10 and 17 months). See Table 1 for details. Maternal education level was used as a proxy for families' socio-economic status. Maternal education levels ranged from 'completed school' and 'a doctoral degree' across labs, with a median of 'university degree'. A Kruskal-Wallis test showed that there were significant differences in maternal education across labs, $H = 62.953$, $df = 12$, $p < .001$. As seen in SI3, the median education level in all laboratories was either 'university degree' or 'master degree', the minimum education level ranged from 'completed school' to 'completed college or technical degree', and the maximum level ranged from 'completed university' to 'a doctoral degree'.

All infants were reported to be typically-developing, born full-term, in good physical health, had no history of severe ear infections or hearing deficits, and were not at familial risk for developmental cognitive or language disorders. No information about infants' race and ethnicity was collected for this study. Infants who did not meet these criteria were excluded prior to recruitment when possible. In addition, data for 242 infants were collected but excluded from the final analyses due to failure to comply with all the inclusion criteria (listed in detail in SI4). All caregivers provided written consent prior to participating in the study (see SI2 for detailed ethics information).

Monolingual infants. Infants ($N = 364$) were included in one of the monolingual groups in this study if they received at least 90% of weekly exposure to their native language. Monolingual infants were exposed to Cantonese, Dutch, English, French, German, Mandarin, Norwegian, and Swedish. Of these infants, 92 were reported to receive exposure to an additional language for less than 10% of a child's weekly awake time ($M = 92.33\%$, $SD = 2.65$), and it was ensured that this was not a tone or a pitch-accent language.

Bilingual infants. Infants ($N = 98$) were included in one of the bilingual groups in this study if they received a maximum of 75% and a minimum of 25% to each of their

TONE DISCRIMINATION IN INFANCY

languages (Pearson, Fernandez, Lewedeg, & Oller, 1997). Bilingual infants were exposed to the following language pairs: Basque-Spanish, Mandarin-English (infants could be dominant in either of the languages in each pair). Infants' exposure to their dominant language ranged from 48% to 78% ($M = 62.37\%$, $SD = 7.96$), and to their non-dominant language ranged from 22% to 50% ($M = 37.01\%$, $SD = 7.76$). In addition, 24 infants were reported to receive no more than 10% of exposure to a third language ($M = 3.37\%$, $SD = 2.32$). Detailed information about bilingual infants' language exposure patterns is reported in SI5.

Language background and demographic information.

Language background questionnaire. Information about infants' language exposure was collected using the LEAT: Language Exposure Assessment Tool (DeAnda, Bosch, Poulin-Dubois, Zesiger, & Friend, 2016). The LEAT was chosen for this study as it is a parental questionnaire that has been successfully validated across two bilingual populations from different language backgrounds (English-French and English-Spanish), and it is accompanied by a published user manual, which allowed for standardisation of administration and scoring across the labs in this study. The LEAT collects information about the infant's primary interlocutors (who interact with the infant on a weekly basis), the languages that they use, the number of hours that they spend using these languages around the infant every day, and the changes in these patterns of exposure across the infant's lifetime.

Experimental Setup and Apparatus

The specific characteristics of the laboratory space, furniture, hardware, and software used for data collection varied across laboratories. Detailed descriptions of each lab's setup can be found at osf.io/t6sp2. All labs conducted the testing sessions in a sound-attenuated or a quiet room inside an infant laboratory. The testing space was equipped with a computer monitor or a TV screen, centrally-positioned loudspeakers hidden behind or underneath the

TONE DISCRIMINATION IN INFANCY

screen, and a CCTV camera or a webcam located on top of or under the screen focused on the infant's face. The camera was connected to a monitor inside a different room (control room) where the experimenter observed the infant's behavior during the task and coded the infant's responses. The experimental paradigm was run using Habit X software (1000, 2, or Windows versions; Oakes, Sperka, DeBolt, & Cantrell, 2019) installed on a Mac or a Windows computer.

Stimuli

Audio stimuli. A female native Cantonese speaker from Hong Kong was recorded producing instances of the pseudoword /lɪn/ with Cantonese tones 25 (rising tone), 21 (falling tone), and 33 (mid-level tone). The pseudoword was chosen to ensure that it was an acceptable lexical form for infants of all language backgrounds involved in the study, and that it was either not a real word or was a word that would be unlikely to be known to young infants in all these languages (/lɪ:n 25/ = “necklace/to grasp” in Cantonese, /lɪ:n 35/= surname in Mandarin, /lɪ:n/ = female first name in English, Norwegian and Swedish). The tones were chosen for two reasons. First, neither tone 33 nor 21 have corresponding tones in the Mandarin tone inventory ensuring that these tones would be non-native for Mandarin infants learning a tone language. Second, tone 33 was categorised at near-perfect levels by tone and non-tone language speaking adults (Francis, Ciocca, Ma, & Fenn, 2008). Categorisation of tone 21 was lower for non-native speakers and uniformly low across Mandarin and English speakers. In addition, we sought to include both a dynamic and static comparison tone on account of past studies suggesting that infants' sensitivity to tone contour may be influential in non-native and native tone sensitivity (Burnham, Singh, Mattock, Woo, & Kalashnikova, 2018).

TONE DISCRIMINATION IN INFANCY

The speaker was instructed to produce the stimuli in a lively child-directed manner. Four tokens of each tone contour were selected as the final stimuli. Details about the acoustic characteristics of the stimuli are presented in Table 2.

The four tokens for each tone-bearing syllable were concatenated into 20-second strings (each including 14 tokens of the stimuli) with ISIs ranging from 800 to 1000 msec. The stimuli string used for habituation, re-familiarisation, and habituated test trials contained the word with tone 25. The stimuli strings used for the novel test trials contained the word with tone 25 alternating with the word with tone 21 for the perceptually similar trials (e.g., /lɪ:n25/, /lɪ:n21/, /lɪ:n25/, /lɪ:n21/, ...) and alternating with tone 33 for the perceptually distinct trials (e.g., /lɪ:n25/, /lɪ:n33/, /lɪ:n25/, /lɪ:n33/, ...). An additional string with the repetitions of the non-word /pa:k/ was used for the pre- and post-test trials. A musical string was used to accompany the attention-getter. The volume of all the auditory stimuli was set to 65dB. Labs were instructed to set the speaker volume to the level that they customarily use in their studies, and to keep the same level for infants from all age groups in this study.

Visual stimuli. A single static image of a colorful bullseye on a white background was used during the habituation, re-familiarisation, and test trials. An image of the same bullseye retracting and expanding on the screen was used for the attention-getter between trials, and a video of a moving waterwheel toy presented on a black background was used for the pre- and post-test trials. All auditory and visual stimuli used in the experiment can be accessed at osf.io/t6sp2.

Experimental Paradigm

An adaptation of the infant-controlled alternating stimulus paradigm was used (Tyler, Best, Goldstein, & Antoniou, 2014), which allows a within-subjects measure of discrimination of two sound contrasts and previously applied to studies of infant tone perception (Singh et al., 2018). This is particularly relevant for bilingual samples where the

TONE DISCRIMINATION IN INFANCY

degree of bilingualism is difficult to equalise across samples, which motivated our choice of a within-subjects design. This paradigm involves four phases: habituation, test block 1, re-familiarisation, and test block 2. Each test block consisted of two test trials, one habituated trial in which infants were presented with repetitions of the habituated stimulus, and one novel trial in which infants were presented with the habituated stimulus alternating with a novel stimulus. The alternating stimuli were different in each test block (e.g., the perceptually similar contrast in Block 1, and the perceptually distinct contrast in Block 2, counterbalanced across infants). The re-familiarisation phase, which consisted of three trials of the habituation stimulus, was presented between the two test blocks. This phase served the purpose of reinstating the habituated stimulus as the baseline before presenting infants with a new test stimulus. In order to avoid losing the infants' attention during re-familiarisation, and to maintain the overall brief duration of the task, no habituation criterion was imposed in this phase.

The task started by presenting the infant with the attention-getter stimulus. Infants also saw the same attention-getter before the start of every trial. After the infant fixated the center of the screen, they saw the pre-test trial. Next, the habituation phase began. This phase continued until infants reached the habituation criterion of a mean 50% decrease in looking time in three consecutive trials compared to the baseline calculated as the average looking time in the first three habituation trials. If infants failed to reach this criterion, the habituation phase continued up to a maximum of 24 trials, and these infants' data were excluded from the final analyses.

After habituation, infants proceeded to complete the first test block, three re-familiarisation trials, and the second test block. Each test block consisted of one habituated trial and one novel trial, and the re-familiarisation trials consisted of the same stimuli used in the habituation phase. One test block included the similar contrast test trial (25-21), and the

TONE DISCRIMINATION IN INFANCY

other block the distinct contrast test trial (25-33). The order of presentation of the two test blocks (similar first vs. distinct first) was counterbalanced across participants within each age and language group in each lab. The trial order was also counterbalanced across blocks within participants (i.e., novel trial presented first in Block 1 and habituated trial presented first in Block 2 and vice versa). At the end of the task, infants were presented with the post-test trial.

The duration of all trials (habituation, re-familiarisation, test, pre- and post-test) was infant-controlled with a maximum length of 20 seconds, a minimum length of 1 second, with the end of the trial triggered by a look away time of > 1 second.

During the task, infants sat in a dimly lit infant laboratory room on their caregiver's lap facing the experimental display. Caregivers listened to masking sounds over headphones (a concatenation of music and the experimental stimuli; Nelson et al., 1995) and were instructed to avoid speaking to the infant and/or pointing to the screen.

Results

Linear Mixed Effects (LME) models were used for all analyses included in this study. Analyses were conducted using the `lme4` package (Bates, 2005) in R (R Core Team, 2013) and the `lmerTest` package (Kuznetsova, Brockhoff, & Christensen, 2015) was used to compute *p*-values. The Independent Variables defined for all models were: Age in weeks (centered before entering the models), Language Background (monolingual infants, bilingual infants), Language Type (tone language, pitch-accent language, non-tone language), Contrast Type (distinct, similar), and Trial (habituated, novel). The dependent variable in all LME models was infants' looking time across different trials. Random intercepts were specified for participant and laboratory. Our analyses did not aim to compare infant performance across individual laboratories, but a figure of infants' performance by laboratory and age group is

presented in SI6. Two sets of analyses were conducted: Confirmatory Analyses and Exploratory Analyses. Detailed rationale for the analyses and for each LME model's specification is provided in SI7.

All the confirmatory and exploratory analyses reported below focused on infants' performance in the test phase. In order to get a deeper understanding of infants' performance patterns in this task, within and across individuals, we also conducted exploratory analyses of infants' performance in each phase of the task (habituation, re-familiarisation, and test). Results and visualisations of these analyses are reported in SI8.

Confirmatory Analyses

Model 1: Effects of Age, Tone Contrast Type, and Language Background on Infants' Non-Native Tone Discrimination

LME Model 1 assessed the effects of Age (5 to 17 months), Tone Contrast Type (similar, distinct) and Language Background (monolingual, bilingual) on infants' tone discrimination (Figure 3). As this analysis focused on non-native tone discrimination, the subset of infants who were not acquiring Cantonese ($N = 416$) was included. The results are presented in Table 3 (see SI9 for model fit). A main effect of Trial indicated that infants looked significantly longer in response to novel compared to habituated trials. There was also a main effect of Language Background; bilingual infants' looking times were significantly shorter compared to monolingual infants. Finally, a marginal effect of Tone Contrast Type suggested that infants' looking times were longer in the distinct versus similar tone contrast test block. All other main effects and interactions were not statistically significant.

Model 2: Effects of Age, Tone Contrast Type, and Language Type on Infants' Non-Native Tone Discrimination

LME Model 2 assessed the effects of Age (5 to 17 months), Tone Contrast Type (similar, distinct) and Language Type (tone, non-tone, pitch-accent) on infants' tone

TONE DISCRIMINATION IN INFANCY

discrimination (Figure 4). This analysis also only included the subset of infants who were not acquiring Cantonese ($N = 416$). The results are presented in Table 4 (see SI9 for model fit). A main effect of Trial indicated that infants looked significantly longer in response to novel compared to habituated trials. There was also a marginal effect of Tone Contrast Type suggesting that infants' looking times were longer in the distinct versus similar tone contrast test block. All other main effects and interactions were not statistically significant.

Model 3: Effects of Tone Language Exposure (Native vs. Non-native Cantonese Tones), Age, Tone Contrast Type, and Language Background on Tone Language-learning Infants' Tone Discrimination

LME Model 3 assessed the effects of infants' exposure to Cantonese tones (native, non-native), Age (5 to 17 months), Tone Contrast Type (similar, distinct), and Language Background (monolingual, bilingual) on infants' tone discrimination (Figure 5). Since Cantonese tones were used as stimuli in this study, the native tone group included monolingual Cantonese infants, and the non-native tone group included monolingual and bilingual infants acquiring Mandarin ($N = 157$). Results are presented in Table 5 (see SI9 for model fit). This model did not yield any significant main effects or interactions.

Exploratory Analyses: Effects of Tone Exposure on Bilingual Infants' Non-Native Tone Discrimination

One exploratory analysis was pre-registered to assess the effects of bilingual infants' individual language experience on tone discrimination. This analysis was limited to the bilingual infants ($N = 98$), and included the additional factor Tone Exposure. Tone Exposure reflected infants' language dominance and whether they were acquiring a tone/non-tone or non-tone/non-tone language pair. This continuous variable was computed as the percentage of exposure to a tone language by a bilingual child (ranging from 0 to 75%, where a non-tone/non-tone bilingual would have a score of 0, and a tone/non-tone bilingual with 75%

exposure to the tone language, the maximum possible given our inclusion criteria, would have a score of 75). This analysis explored how this factor related to bilingual exposure interacts with Age and Trial factors included in our research design. Results are presented in Table 6 (see SI9 for model fit). A main effect of Trial indicated that overall, bilingual infants looked longer in response to novel than habituated trials, but there were no other significant main effects or interactions (see Figure 6).

Finally, in a set of *post-hoc* exploratory analyses that were not part of the original pre-registration, we assessed the possibility that infants' non-native and native tone discrimination performance in our task could be affected by the order in which test blocks were presented during the test phase. These analyses are reported in SI10.

Discussion

The purpose of this study was to investigate infants' discrimination of Cantonese tones in relation to a number of factors, specifically, infant age, tone language experience, bilingual experience, contrast type, and exposure to Cantonese tones (native vs. non-native discrimination). We set out to test a series of predictions aimed at investigating prior mixed findings on age-related change in infant tone discrimination. Testing a large, diverse sample of infants, we first examined effects of age to determine whether infants demonstrate age-related decline, facilitation, or maintenance in tone discrimination, all trajectories that have been reported in prior infant tone discrimination research (Figure 2). Our findings indicate robust sensitivity to two Cantonese tone contrasts in infants acquiring languages different from Cantonese. Further, this sensitivity does not appear to be modulated by age, at least not from 5 to 17 months. This points to maintenance of tone sensitivity over the age groups sampled.

Traditional models of perceptual reorganisation based on consonant and vowel perception would predict interactions of age and native language experience in relation to the

TONE DISCRIMINATION IN INFANCY

contrasts used (Werker, 2018). To test this prediction, we examined whether two different types of language experience impacted tone sensitivity. The first was whether or not pitch is used at the lexical level in infants' native language. The second was whether infants had native familiarity with the Cantonese tones used (i.e., whether Cantonese was one of their native language(s)) or whether they were learning a different tone language (i.e., Mandarin). Despite past evidence demonstrating experience-dependent effects on infant tone discrimination, our analyses did not reveal variation in tone discrimination based on either of these two forms of experience.

Our study also examined the effects of bilingualism on the developmental trajectory of infants' tone discrimination. Bilingual infants in our sample showed comparable discrimination patterns to their monolingual peers, suggesting maintenance of their discrimination ability from 5 to 17 months. This pattern differs from previous studies with infants acquiring two non-tone languages, which have shown a U-shaped function whereby infants demonstrate a temporary decline followed by resurgence in non-native tone discrimination (Götz et al., 2018; Liu & Kager, 2014; 2017). It is challenging to draw direct comparisons across studies since our sample combined infants acquiring two non-tone languages (Basque-Spanish) and one tone and one non-tone language (Mandarin-English). Our planned exploratory analyses suggest that infants' performance was not impacted by the varying degree of exposure to a tone language, but further studies of diverse bilingual populations could elucidate these contradictory findings. It is also noteworthy that despite our efforts to include different linguistic profiles in our sample, we were unable to recruit bilingual infants acquiring two tone languages (i.e., Cantonese-Mandarin), a population that is critical for a complete understanding of the effects of diverse language experience on infants' sensitivity to native and non-native lexical tone categories.

TONE DISCRIMINATION IN INFANCY

An unexpected finding was that the subset of infants who were learning a tone language (either Cantonese or Mandarin) did not show evidence of discriminating the Cantonese contrasts used. Unlike our main analysis model that included the entire sample, this model included only a smaller subset of tone language learning infants (N=157). However, the sample size far exceeded that of prior studies demonstrating successful tone discrimination in native learners of tone languages. For example, Yeung and colleagues (2013) demonstrated that both Mandarin- and Cantonese-exposed infants at 4 to 9 months of age discriminated the same distinct tone contrast used in the present study in a much smaller sample. One possibility for the difference in findings is that Yeung et al. (2013) focused exclusively on monolingual learners and on a narrower age range (4 to 9 months). The presence of a more varied sample here – both in terms of age and language background – may have attenuated evidence of discrimination. In addition, there were differences in the paradigm used by Yeung et al. (2013) and this study. In general, studies on perceptual narrowing have been shown to be susceptible to subtle variation in paradigms (see Singh, Rajendra, & Mazuka, 2022 for a discussion of this issue). The discrepancy in findings and absence of discrimination in tone language learners motivates greater replication efforts to determine the reliability of reported effects.

Finally, there was no effect of contrast type on lexical tone discrimination or an interaction of contrast type with age. Effects of similar versus distinct contrasts on infant tone discrimination have been demonstrated in past studies, with distinct contrasts associated with clear discrimination advantages (e.g., Liu & Kager, 2014; Singh et al., 2018; Tsao, 2017). The present study suggests that tone properties did not predict discrimination, nor did this factor interact with age and tone language experience. These results are consistent with a recent study demonstrating that the neural encoding of lexical tones develops rapidly in the first two years of life for tone-learning children regardless of the nativeness and

TONE DISCRIMINATION IN INFANCY

psychoacoustic properties of the tone (Novitsky et al., 2022). Furthermore, neural responses to both native and non-native tones in infancy predicted language development in toddlerhood (Wong et al., 2021). Nonetheless, it is possible that the two Cantonese contrasts tested here were perceptually salient for infants, and that a contrast effect may become apparent with contrasts from tone languages other than Cantonese. This possibility is suggested by a recent comprehensive study of the psychoacoustic properties of tones on adult tone discrimination (Liu et al., 2022), but this evidence is not yet available for infants.

Against a backdrop of theories about how infants break into native language structure, the greatest theoretical significance of the current findings was the absence of any evidence for age-related decline in tone language sensitivity and of any interaction of age and tone language experience (see SI11 for a depiction of developmental trajectories in looking times to habituated and novel trials for infants acquiring each language type). Theories of perceptual narrowing rely crucially on an account of perceptual development where, over time, infants' perceptual sensitivities shift in alignment with experience with the stimuli used (Gervain, 2022). Age-related change in sensitivity to both native and non-native contrasts are central postulates of this influential theory (but see Best, 1994). In line with past studies that have cast doubt on whether this interaction applies to lexical tones, the present findings suggest that the developmental trajectory associated with tone discrimination may vary from the trajectory commonly charted for vowels and consonants. If anything, the absence of tone discrimination in native learners of tone languages in the age range tested here does not fit a central premise of perceptual narrowing theory.

In considering why the developmental trajectory associated with tone discrimination varies from prior studies focusing on vowels and consonants, we suggest several reasons for a deviation with respect to lexical tone. First, tone is carried by pitch, which in turn conveys a broad range of communicative functions (Yuan, 2011). In every language, pitch movements

TONE DISCRIMINATION IN INFANCY

are relevant to linguistic communication. For example, pitch is used to signal intonational contrast, communicative intent, emotion, and stress (see Bolinger, 1978; Ladefoged & Disner, 2012; Lieberman, 1967). Therefore, learners benefit from attending to pitch movements even if these are not lexically contrastive. Studies investigating the earliest phases of word recognition have demonstrated that infants closely attend to pitch changes across words, equating words that are matched in pitch whether they are learning a tone language or not (Singh & Foong, 2012). These findings suggest that the universal use of pitch movements in the service of communication may bias early perceptual systems towards pitch sensitivity (see Hay et al., 2019). This may not be the case for vowels and consonants and particularly for those that have been used in past studies of perceptual narrowing. While there is variation in the proximity between vowel and consonant stimuli used in past studies and infants' auditory experiences (Best, 1995), in general, vowels and consonants that provide the evidence basis for theories of perceptual narrowing are either lexically contrastive or not in infants' native languages. For example, the most commonly used contrast to demonstrate age-related decline is the Hindi dental-retroflex contrast in English learners (Singh et al., 2022). This particular contrast is likely entirely absent from the linguistic interactions of most English-learning infants.

Pitch movements, in contrast, are ubiquitous in human language. The types of pitch movements that drive changes in lexical tone often overlap with intonational categories (Braun & Johnson, 2011; Wang, Kager, & Wong, 2022; Wang et al., 2021). It is possible that the disaggregation of pitch into lexical and non-lexical function is inherently more complex than the differentiation of consonants and vowels as native or non-native (see Kager, 2018 for a discussion of this issue). Due to increased complexity, learners may retain tone sensitivity in the service of disaggregating pitch into its constituent functions. Some support for this notion comes from studies investigating tone sensitivity in word learners that establish that

even English learners integrate tone information into newly learned words (e.g., Singh et al., 2014). In particular, tones that are integrated into newly-learned words are those that overlap with native intonational categories (Hay et al., 2019). Furthermore, the disaggregation of pitch movements into tonal and intonational categories in word recognition matures relatively late during the preschool years in tone language learners (Singh & Chee, 2016). These findings suggest a functional complexity account where phonological units that serve multiple communicative functions may be navigated along an extended timetable than those that primarily serve to draw lexical contrast.

A second possibility for why the current study failed to demonstrate developmental patterns characteristic of perceptual narrowing theory relates to the evidence basis for perceptual narrowing studies thus far. A recent meta-analysis suggests that although perceptual narrowing is expressed as a universal phenomenon (Kuhl, 2004; Werker & Tees, 2002), the empirical evidence that has informed the theory is far from universal (Singh et al., 2022). Relying heavily on discrimination of consonants and vowels, the populations are heavily skewed towards North America and Western Europe. Additionally, the stimuli used constitute a very limited set of sounds that appear in a small number of languages of the world. For native and non-native sounds, there is strong representation of particular Indo-European languages. Asia, the world's most populous continent, has received much less attention in the perceptual narrowing literature. This disproportionate focus on particular world regions has had two effects relevant to this study. First, it has led to very marginal treatment of tones, which are most represented in African and Asian countries. Second, the literature has focused heavily on particular populations raising questions about generalisability. Recent studies on perceptual narrowing in under-studied populations show that even for vowels and consonants, the typical pattern of perceptual narrowing is not always upheld (for Japanese, see Sato et al., 2010; 2012; for Korean, Shin et al., 2018). It is therefore

possible that this sampling bias that has led to overrepresentation of particular language communities, phonetic contrasts, and geographical regions has constructed a strong narrative of perceptual narrowing and that as a theory, it may not be as generalisable as originally proposed. As a greater diversity of evidence is collected, cumulative data will help to establish the robustness of the theory and identify boundary conditions for reported effects.

In terms of limitations and constraints on generalisability, our study used a single paradigm to measure developmental change in tone perception. In addition, we used as stimuli two pairs of tones from a single language. Studies on infant speech discrimination have been shown to be highly vulnerable to methodological choices. Seemingly minor changes (e.g., order of stimulus presentation, inter-stimulus interval, task demands) can markedly shift the expression of discrimination (e.g., Werker & Logan, 1985). For example, Yeung et al. (2014) found that English-learning infants' sensitivity to the distinct pair of Cantonese tones used in our study was higher when the tones were paired with objects rather than a checkerboard pattern, and that this effect was modulated by vocabulary size. This suggests that experimental modifications have significant effects on tone discrimination. Another methodological choice in our study consisted of using a single habituation stimulus for all infants instead of counterbalancing between the habituation and test stimuli. Capturing potential asymmetries in infants' early tone discrimination was not one of our goals (see Liu et al., 2018; Politzer-Ahles et al., 2016 for reports of perceptual asymmetries), but it is possible that including this manipulation may have induced directional effects of the perceptually similar and distinct tone contrasts on infants' performance. Finally, tone inventories are heterogenous and diverse even within East Asian languages (Gandour, 1983). Past studies have shown that the use of different tones and different tone languages elicit different sensitivities in infants (e.g., Burnham et al., 2018; Feng, Kager, Lai, & Wong, 2022; Hay et al., 2019) and adults (Liu et al., 2022). It is possible that the current set of findings

TONE DISCRIMINATION IN INFANCY

applies to the specific tone contrasts used and while testing with infants precludes the use of multiple contrasts, interpretation of findings is constrained by the use of two tone pairs from a single language.

To summarise, the present study expands and diversifies the narrative around perceptual reorganisation by studying sensitivity to a property of speech, lexical tone, that is widely represented across languages of the world. In addition, we have sampled broadly across a range of societies, languages, and types of learners, providing greater participant heterogeneity than a typical single-site laboratory study. Adopting this multi-laboratory approach allows us to examine crucial issues of generalisability and stability of a widely-documented developmental transition. Our findings suggest that discrimination of lexical tones may be associated with maintenance between 5 and 17 months, aligning neither with attenuation or facilitation. These findings are a critical addition to the evidence underlying fundamental theories of development, which must account for the full diversity of learners and languages around the world.

References

- Aoyama, K., Flege, J. E., Guinon, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /ɾ/. *Journal of Phonetics*, 32, 233-250.
- Aslin, R. N., & Pisoni, D. B. (1980). Some developmental processes in speech perception. In G. H. Yeni-Komshian, J. F. Kavanagh & C. A. Ferguson (Eds.), *Child Phonology* (Vol. 2: Perception). New York: Academic Press.
- Bates, D. (2005). Fitting linear mixed models in R. *R News*, 5(1), 27–30.
- Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. *The development of speech perception: The transition from speech sounds to spoken words*, 167(224), 233-277.
- Best, C. T. (1995). Learning to perceive the sound pattern of English. *Advances in infancy research*, 9, 217-217.
- Bolinger, D. (1978). Intonation across languages. *Universals of human language*, II: Phonology, ed. by Joseph H. Greenberg et al., 471-524.
- Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, 46(2-3), 217-243.
- Braun, B., & Johnson, E. K. (2011). Question or tone 2? How language experience and linguistic function guide pitch processing. *Journal of Phonetics*, 39(4), 585-594.
- Burnham, D. K. (1986). Developmental loss of speech perception: Exposure to and experience with a first language. *Applied Psycholinguistics*, 7(3), 207-239.
- Burnham, D., Kasisopa, B., Reid, A., Luksaneeyanawin, S., Lacerda, F., Attina, V., ... & Webster, D. (2015). Universality and language-specific experience in the perception of lexical tone and pitch. *Applied Psycholinguistics*, 36(6), 1459-1491.

TONE DISCRIMINATION IN INFANCY

- Burnham, D., Singh, L., Mattock, K., Woo, P. J., & Kalashnikova, M. (2018). Constraints on tone sensitivity in novel word learning by monolingual and bilingual infants: tone properties are more influential than tone familiarity. *Frontiers in Psychology, 8*, 2190.
- Byers-Heinlein, K., Tsui, A. S. M., Bergmann, C., Black, A. K., Brown, A., Carbajal, M. J., ... & Wermelinger, S. (2021). A multilab study of bilingual infants: Exploring the preference for infant-directed speech. *Advances in Methods and Practices in Psychological Science, 4*(1), 2515245920974622.
- Cabrera, L., Tsao, F. M., Liu, H. M., Li, L. Y., Hu, Y. H., Lorenzi, C., et al. (2015). The perception of speech modulation cues in lexical tones is guided by early language-specific experience. *Frontiers in Psychology, 6*, 1290. doi: 10.3389/fpsyg.2015.01290
- Chao, Y. R. (1968). *A Grammar of Spoken Chinese*. Berkeley: California University Press.
- Chen, A., & Kager, R. (2016). Discrimination of lexical tones in the first year of life. *Infant and Child Development, 25*(5), 426-439.
- Chen, H., Lee, D.T., Luo, Z., Lai, R.Y., Cheung, H., & Nazzi, T. (2021). Variation in phonological bias: Bias for vowels, rather than consonants or tones in lexical processing by Cantonese-learning toddlers. *Cognition, 213*, 104486. DOI: 10.1016/j.cognition.2020.104486
- Chen, A., Stevens, C. J., & Kager, R. (2017). Pitch perception in the first year of life: A comparison of lexical tones and musical pitch. *Frontiers in Psychology, 8*, 1–8. <https://doi.org/10.3389/fpsyg.2017.00297>
- DeAnda, S., Bosch, L., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2016). The language exposure assessment tool: Quantifying language exposure in infants and children. *Journal of Speech Language and Hearing Research, 59*, 1346–4356.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science, 171*(3968), 303–306.

- Feng, Y., Kager, R., Lai, R., & Wong, P. C. M. (2022). The ability to use contextual cues to achieve phonological constancy emerges by 14 months. *Developmental Psychology*, 58(11), 2064–2080. <https://doi.org/10.1037/dev0001418>
- Ferjan Ramírez, N., Ramírez, R. R., Clarke, M., Taulu, S., & Kuhl, P. K. (2017). Speech discrimination in 11-month-old bilingual and monolingual infants: a magnetoencephalography study. *Developmental Science*, 20(1), e12427.
- Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics*, 36(2), 268-294.
- Frank, M. C., Bergelson, E., Bergmann, C., Cristia, A., Floccia, C., Gervain, J., ... & Lew-Williams, C. (2017). A collaborative approach to infant research: Promoting reproducibility, best practices, and theory-building. *Infancy*, 22(4), 421-435
- Fromkin, V. (ed.) (1978). *Tone: A Linguistic Survey*. New York, NY: Academic Press.
- Gervain, J. (2022). Development of Speech Perception. In *Speech Perception* (pp. 201-226). Springer, Cham.
- Gómez, D. M., Mok, P., Ordin, M., Mehler, J., & Nespors, M. (2018). Statistical speech segmentation in tone languages: The role of lexical tones. *Language and Speech*, 61(1), 84-96.
- Götz, A., Yeung, H. H., Krasotkina, A., Schwarzer, G., & Höhle, B. (2018). Perceptual reorganization of lexical tones: effects of age and experimental procedure. *Frontiers in Psychology*, 9, 477.
- Hay, J. F., Cannistraci, R., & Zhao, Q. (2019). Mapping non-native pitch contours to meaning: Perceptual and experiential factors. *Journal of Memory and Language*.
- Hay, J. F., Graf Estes, K., Wang, T., & Saffran, J. R. (2015). From flexibility to constraint: The contrastive use of lexical tone in early word learning. *Child Development*, 86, 10-

22. <https://doi.org/10.1111/cdev.12269>

Hyman, L. M. (2009). How (not) to do phonological typology: the case of pitch-accent.

Language Sciences, 31(2-3), 213-238.

Kager, R. (2018) How do infants disaggregate referential and affective pitch? *Frontiers in*

Psychology. 9:2093. doi: 10.3389/fpsyg.2018.02093

Klein, R., Ratliff, K., Vianello, M., Adams Jr, R., Bahník, S., Bernstein, M., ... & Cemalcilar,

Z. (2014). Data from investigating variation in replicability: A “many labs” replication project. *Journal of Open Psychology Data*, 2(1).

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews*

Neuroscience, 5(11), 831.

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. doi:10.1111/j.1467-7687.2006.00468.x

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(15), 9096–9101.

<https://doi.org/10.1073/pnas.1532872100>

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. doi:10.1126/science.1736364

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). Package “lmerTest.” *R*

Package Version, 2(0).

Ladefoged, P., & Disner, S. F. (2012). *Vowels and consonants*. John Wiley & Sons.

Ladefoged, P., & Disner, S. F. (2012). *Vowels and consonants*. John Wiley & Sons.

Lieberman, P. (1967). Intonation, perception, and language. *MIT Research Monograph*.

TONE DISCRIMINATION IN INFANCY

- Liu, L., & Kager, R. (2014). Perception of a native vowel contrast by Dutch monolingual and bilingual infants: A bilingual perceptual lead. *International Journal of Bilingualism*, 20(3), 335–345. <https://doi.org/10.1177/1367006914566082>
- Liu, L., & Kager, R. (2016). Perception of a native vowel contrast by Dutch monolingual and bilingual infants: A bilingual perceptual lead. *International Journal of Bilingualism*, 20(3), 335-345.
- Liu, L., & Kager, R. (2017). Perception of tones by bilingual infants learning non-tone languages. *Bilingualism: Language and Cognition*, 20(3), 561-575.
- Liu, L., & Kager, R. (2018). Monolingual and bilingual infants' ability to use non-native tone for word learning deteriorates by the second year after birth. *Frontiers in Psychology*, 9, 117.
- Liu, L., Lai, R., Singh, L., Kalashnikova, M., Wong, P. C., Kasisopa, B., ... & Burnham, D. (2022). The tone atlas of perceptual discriminability and perceptual distance: Four tone languages and five language groups. *Brain and Language*, 229, 105106.
- Liu, L., Ong, J. H., Tuninetti, A., & Escudero, P. (2018). One way or another: Evidence for perceptual asymmetry in pre-attentive learning of non-native contrasts. *Frontiers in Psychology*, 9, 162.
- Liu, S., & Samuel, A. G. (2004). Perception of Mandarin lexical tones when F0 information is neutralized. *Language and speech*, 47(2), 109-138.
- Maddieson, I. (2013). Tone. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/13>)
- ManyBabies Consortium (accepted pending data collection). Quantifying sources of variability in infancy research using the infant-directed speech preference. *Advances in Methods and Practices in Psychological Science*.

TONE DISCRIMINATION IN INFANCY

- Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy, 10*(3), 241–265.
- Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition, 106*(3), 1367–1381.
<https://doi.org/10.1016/j.cognition.2007.07.002>
- Mazuka, R., Hasegawa, M., & Tsuji, S. (2014). Development of non-native vowel discrimination: Improvement without exposure. *Developmental Psychobiology, 56*(2), 192–209. doi: 10.1002/dev.21193
- Narayan, C. R., Werker, J. F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception : Evidence from nasal place discrimination. *Developmental Science, 13*(3), 407–420.
<https://doi.org/10.1111/j.1467-7687.2009.00898.x>
- Nelson, D. G. K., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. (1995). The head-turn preference for testing auditory perception. *Infant Behavior and Development, 18*, 111–116.
- Novitskiy, N., Maggu, A. R., Lai, C. M., Chan, P. H., Wong, K. H., Lam, H. S., ... & Wong, P. C. (2022). Early development of neural speech encoding depends on age but not native language status: Evidence from lexical tone. *Neurobiology of Language, 3*(1), 67-86.
- Pearson, B. Z., Fernandez, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics, 18*(01), 41.
<https://doi.org/10.1017/S0142716400009863>
- Petitto, L. A., Berens, M. S., Kovelman, I., Dubins, M. H., Jasinska, K., & Shalinsky, M. (2012). The “Perceptual Wedge Hypothesis” as the basis for bilingual babies’ phonetic

TONE DISCRIMINATION IN INFANCY

- processing advantage: New insights from fNIRS brain imaging. *Brain and Language*, *121*(2), 130-143.
- Politzer-Ahles, S., Schluter, K., Wu, K., & Almeida, D. (2016). Asymmetries in the perception of Mandarin tones: Evidence from mismatch negativity. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(10), 1547.
- Polka, L., & Bohn, O. S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, *39*(4), 467–478.
<https://doi.org/10.1016/j.wocn.2010.08.007>
- Polka, L., & Werker, J. F. (1994). Developmental Changes in Perception of Nonnative Vowel Contrasts. *Journal of Experimental Psychology Human Perception & Performance*, *20*(2), 421–435.
- Poltrock, S., Chen, H., Kwok, C., Cheung, H., & Nazzi, T. (2018). Adult learning of words in a nonnative language: Consonants, vowels, and tones. *Frontiers in Psychology*, *9*(1211). 10.3389/fpsyg.2018.01211.
- Ramachers, S., Brouwer, S., & Fikkert, P. (2018). No perceptual reorganization for Limburgian tones? A cross-linguistic investigation with 6-to 12-month-old infants. *Journal of Child Language*, *45*(2), 290-318.
- Sato, Y., Kato, M., & Mazuka, R. (2012). Development of single/geminate obstruent discrimination by Japanese infants: Early integration of durational and nondurational cues. *Developmental Psychology*, *48*(1), 18-34
- Sato, Y., Sogabe, Y., & Mazuka, R. (2010). Discrimination of phonemic vowel length by Japanese infants. *Developmental Psychology*, *46*(1), 106-119.
- Shi, R., Santos, E., Gao, J., & Li, A. (2017). Perception of similar and dissimilar lexical tones by non-tone-learning infants. *Infancy*, *22*(6), 790–800.
<https://doi.org/10.1111/infa.12191>

TONE DISCRIMINATION IN INFANCY

- Shin, M., Choi, Y., & Mazuka, R. (2018). Development of fricative sound perception in Korean infants: The role of language experience and infants' initial sensitivity. *PLoS ONE*, *13*(6), e0199045.
- Singh, L., & Chee, M. (2016). Rise and fall: Effects of tone and intonation on spoken word recognition in early childhood. *Journal of Phonetics*, *55*, 109-118.
- Singh, L., & Foong, J. (2012). Influences of lexical tone and pitch on word recognition in bilingual infants. *Cognition*, *124*(2), 128–142.
<https://doi.org/10.1016/j.cognition.2012.05.008>
- Singh, L., & Fu, C. S. L. (2016). A new view of language development: The acquisition of lexical tone. *Child Development*, *97*(3), 834-854.
- Singh, L., Fu, C. S. L., Seet, X. H., Tong, A. P. Y., Wang, J. L., & Best, C. T. (2018). Developmental change in tone perception in Mandarin monolingual, English monolingual, and Mandarin–English bilingual infants: Divergences between monolingual and bilingual learners. *Journal of Experimental Child Psychology*, *173*, 59–77. <https://doi.org/10.1016/j.jecp.2018.03.012>
- Singh, L., Hui, T. J., Chan, C., & Golinkoff, R. M. (2014). Influences of vowel and tone variation on emergent word knowledge: A cross-linguistic investigation. *Developmental Science*, *17*(1), 94-109.
- Singh, L., Poh, F. L. S., & Fu, C. S. L. (2016). Limits on monolingualism? A comparison of monolingual and bilingual infants' abilities to integrate lexical tone in novel word learning. *Frontiers in Psychology*, *7*, 1–16. <https://doi.org/10.3389/fpsyg.2016.00667>
- Singh, L., Rajendra, S. J., & Mazuka, R. (2022). Diversity and representation in studies of infant perceptual narrowing. *Child Development Perspectives*, *16*(4), 191-199.

TONE DISCRIMINATION IN INFANCY

- So, C. K., & Best, C. T. (2010). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech*, *53*(2), 273-293.
- Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*(6640), 381–382.
<https://doi.org/10.1038/41102>
- Tsao, F. (2017). Perceptual improvement of lexical tones in infants: Effects of tone language experience. *Frontiers in Psychology*, *8*, 1–14. <https://doi.org/10.3389/fpsyg.2017.00558>
- Tsao, F. M. (2008). The effect of acoustical similarity on lexical-tone perception of one-year-old Mandarin-learning infants. *Chinese Journal of Psychology*, *50*(2), 111-124.
- Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *The Journal of the Acoustical Society of America*, *120*(4), 2285-2294.
- Tyler, M. D., Best, C. T., Goldstein, L. M., & Antoniou, M. (2014). Investigating the role of articulatory organs and perceptual assimilation in infants' discrimination of native and non-native fricative place contrasts. *Developmental Psychobiology*, *56*(2), 210-227.
- Wang, L. C., Kager, R., & Wong, P. C. M. (2022). The Effect of Tone Hyperarticulation in Cantonese Infant-Directed Speech on Toddlers' Word Recognition in the Second Year of Life. *First Language*, *42*, 670-692, DOI: doi.org/10.1177/01427237221109342
- Wang, L. C., Kalashnikova, M., Kager, R., Lai, R., & Wong, P. C. M. (2021). Lexical and Prosodic Pitch Modifications in Cantonese Infant-directed Speech. *Journal of Child Language*, *48*, 1235–1261, DOI: doi.org/10.1017/S0305000920000707
- Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *The Journal of the Acoustical Society of America*, *106*(6), 3649-3658.

TONE DISCRIMINATION IN INFANCY

- Werker, J. F. (2018). Speech perception, word learning, and language acquisition in infancy: The voyage continues. *Applied Psycholinguistics*, *39*(4), 769-777.
- Werker, J. F., & Hensch, T. K. (2015). Critical periods in speech perception: New directions. *Annual Review of Psychology*, *66*, 173–196. <https://doi.org/10.1146/annurev-psych-010814-015104>
- Werker, J. F., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics*, *37*(1), 35-44.
- Werker, J. F., & Tees, R. C. (1984a). Cross-language speech-perception: Evidence for perceptual reorganization during the 1st year of life. *Infant Behavior and Development*, *7*(1), 49–63. [https://doi.org/Doi.10.1016/S0163-6383\(84\)80022-3](https://doi.org/Doi.10.1016/S0163-6383(84)80022-3)
- Werker, J. F., & Tees, R. C. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *The Journal of the Acoustical Society of America*, *75*(6), 1866-1878.
- Werker, J. F., & Tees, R. C. (2002). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *25*(1), 121-133.
- Wewalaarachchi, T. D., Wong, L. H., & Singh, L. (2017). Vowels, consonants, and lexical tones: Sensitivity to phonological variation in monolingual Mandarin and bilingual English–Mandarin toddlers. *Journal of Experimental Child Psychology*, *159*, 16-33.
- Wiener, S., & Turnbull, R. (2016). Constraints of tones, vowels and consonants on lexical selection in Mandarin Chinese. *Language and Speech*, *59*(1), 59-82.
- Wong, P. C., Lai, C. M., Chan, P. H., Leung, T. F., Lam, H. S., Feng, G., ... & Novitskiy, N. (2021). Neural speech encoding in infancy predicts future language and communication difficulties. *American Journal of Speech-Language Pathology*, *30*(5), 2241-2250.
- Wright, D. B. (2017). Some limits using random slope models to measure academic growth. *Frontiers in Education*, *2* (58). doi: 10.3389/educ.2017.00058

TONE DISCRIMINATION IN INFANCY

Yeung, H. H., Chen, K. H., & Werker, J. F. (2013). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and*

Language, 68(2), 123–139. <https://doi.org/10.1016/j.jml.2012.09.004>

Yeung, H. H., Chen, L. M., & Werker, J. F. (2014). Referential labeling can facilitate phonetic learning in infancy. *Child Development*, 85(3), 1036-1049.

Yip, M. (2002). “Chap. 1, Introduction,” in *Tone* (New York, NY: Cambridge University Press), 1–14.

Yuan, J. (2011). Perception of intonation in Mandarin Chinese. *The Journal of the Acoustical Society of America*, 130(6), 4063-4069.

TONE DISCRIMINATION IN INFANCY

Table 1. *Detailed information about the language background, age, and sex of infants included in each participant group.*

	Lab location	Specific language(s)	<i>N</i>	Age (days)	Sex (M, F)
Tone (Native exposure to Cantonese) monolingual	Hong Kong	Cantonese	46	Range = 131 to 538, M = 345.83, SD = 143.20	18, 28
Tone monolingual	Malaysia, Taiwan, Singapore	Mandarin	56	Range = 120 to 538, M = 323.04, SD = 150.54	30, 26
Non-tone monolingual	Australia, France, Germany, Netherlands, USA	English, French, German, Dutch	209	Range = 101 to 540, M = 309.31, SD = 144.24	112, 97
Pitch-accent monolingual	Norway, Sweden	Norwegian, Swedish	53	Range = 126 to 547, M = 277.89, SD = 130.12	30, 23
Tone/Non-tone bilingual	Malaysia, Singapore	Mandarin and English	54	Range = 133 to 540, M = 362.55, SD = 147.23	29, 25
Non-tone/Non-tone bilingual	Spain	Spanish and Basque	44	Range = 122 to 538, M = 320.70, SD = 157.68	18, 26

TONE DISCRIMINATION IN INFANCY

Table 2. *Acoustic correlates of the stimuli used in this study.*

Tone	Token	Mean (Hz)	Onset (Hz)	Offset (Hz)	Dur (msec)
Tone 25	1	251.980	219.173	298.247	0.510
	2	243.450	211.292	284.359	0.518
	3	257.490	213.294	304.732	0.439
	4	270.019	235.269	305.267	0.482
Tone 21	1	195.847	217.948	189.805	0.524
	2	193.799	226.896	181.154	0.567
	3	194.603	222.434	183.365	0.545
	4	202.482	233.640	184.169	0.473
Tone 33	1	277.402	275.554	276.531	0.633
	2	272.972	267.808	270.244	0.516
	3	272.173	269.120	291.201	0.529
	4	276.377	281.351	247.663	0.452

STONE DISCRIMINATION IN INFANCY

Table 3. Detailed output of Linear Mixed Effects Model 1^a assessing the effects of Trial Type, Age, Contrast Type, and Language Background on infants' non-native tone discrimination (N = 416).

	Estimate	Std. Error	df	t	p
(Intercept)	-8.94E-02	4.79E-02	1.11E+01	-1.864	.089
Age	4.12E-03	3.27E-02	4.05E+02	0.126	.900
Trial (habituated)	-8.91E-02	2.44E-02	1.24E+03	-3.653	<.001
Contrast Type (distinct)	4.03E-02	2.08E-02	1.25E+03	1.939	.053
Language Background (bilingual)	-1.57E-01	4.76E-02	1.24E+01	-3.294	.006
Age × Trial (habituated)	-3.25E-02	2.08E-02	1.24E+03	-1.562	.119
Trial (habituated) × Contrast Type (distinct)	-9.37E-04	2.08E-02	1.24E+03	-0.045	.964
Trial (habituated) × Language Background (bilingual)	6.66E-04	2.44E-02	1.24E+03	0.027	.978

^a Looking time ~ Age*Trial + Contrast Type*Trial + Language Background *Trial + (1 | Lab) + (1 | Subject)

TONE DISCRIMINATION IN INFANCY

Table 4. Detailed output of Linear Mixed Effects Model 2^a assessing the effects of Trial Type, Age, Contrast Type, and Language Type on infants' non-native tone discrimination (N = 416).

	Estimate	Std. Error	df	t	p
(Intercept)	-5.99E-02	6.53E-02	1.04E+01	-0.918	.379
Age	7.53E-03	4.41E-02	3.74E+02	0.17	.865
Trial (habituated)	-9.52E-02	2.60E-02	1.24E+03	-3.67	<.001
Contrast Type (distinct)	4.03E-02	2.08E-02	1.24E+03	1.936	.053
Language Type (non-tone)	1.22E-01	7.93E-02	9.43E+00	1.542	.156
Language type (pitch-accent)	-7.81E-02	1.07E-01	1.09E+01	-0.732	.480
Age × Trial (habituated)	-3.07E-02	2.71E-02	1.24E+03	-1.134	.257
Trial (habituated) × Contrast Type (distinct)	-8.62E-04	2.08E-02	1.24E+03	-0.041	.967
Trial (habituated) × Language Type (non-tone)	1.36E-02	3.02E-02	1.24E+03	0.45	.653
Trial (habituated) × Language Type (pitch-accent)	-1.56E-02	4.37E-02	1.24E+03	-0.357	.721
Age × Language Type (non-tone)	-2.36E-02	5.03E-02	3.91E+02	-0.468	.640
Age × Language Type (pitch-accent)	4.43E-03	7.68E-02	3.59E+02	0.058	.954
Age × Trial (habituated) × Language Type (non-tone)	2.20E-03	3.12E-02	1.24E+03	0.07	.944
Age: Trial (habituated) × Language Type (pitch-accent)	1.67E-02	4.67E-02	1.24E+03	0.357	.721

^a Looking time ~ Age*Trial + Contrast Type*Trial + Language Type*Trial + Age*Language Type*Trial + (1 | Lab) + (1 | Subject)

TONE DISCRIMINATION IN INFANCY

Table 5. Detailed output of Linear Mixed Effects Model 3^a assessing the effects of Trial Type, Age, Contrast Type, and Tone Language Exposure (Native vs. Non-native) on infants' native and non-native tone discrimination (N = 157).

	Estimate	Std. Error	df	t	p
(Intercept)	-0.0796	0.0796	0.9022	-1	.514
Age	0.0510	0.0597	149.6300	0.855	.394
Tone Language [native]	-0.0577	0.0820	0.7285	-0.704	.642
Trial [habituated]	-0.0563	0.0430	464.0055	-1.308	.192
Language Background [bilingual]	-0.1967	0.0730	1.4116	-2.693	.164
Contrast Type [distinct]	0.0031	0.0343	466.4911	0.091	.928
Age × Tone Language [native]	0.0283	0.0597	149.6300	0.475	.636
Age × Trial [habituated]	-0.0322	0.0392	464.0055	-0.821	.412
Tone Language [native]:Trial [habituated]	0.0472	0.0429	464.0055	1.1	.272
Trial [habituated] × Language Background [bilingual]	-0.0220	0.0408	464.0055	-0.54	.589
Trial [habituated] × Contrast Type [distinct]	-0.0145	0.0343	464.0055	-0.422	.673
Age × Tone Language [native] × Trial [habituated]	0.0185	0.0392	464.0055	0.472	.637

^a Looking time ~ Age*Tone Language Exposure*Trial + Language Background*Trial + Contrast Type*Trial + (1 | Lab) + (1 | Subject)

TONE DISCRIMINATION IN INFANCY

Table 6. Detailed output of the exploratory Linear Mixed Effects Model 1^a assessing the effects of Trial Type, Age, and Tone Exposure on infants' native and non-native tone discrimination (N = 98).

	Estimate	Std. Error	df	t	p
(Intercept)	0.006	0.101	1.328	0.062	.959
Age	0.021	0.070	58.175	0.299	.766
Tone Exposure	-0.049	0.097	2.100	-0.508	.660
Trial [habituated]	-0.105	0.044	289.999	-2.377	.018
Age × Tone Exposure	0.062	0.070	67.412	0.889	.377
Age × Trial [habituated]	-0.067	0.044	289.999	-1.529	.127
Tone Exposure × Trial [habituated]	-0.027	0.044	289.999	-0.606	.545
Age × Tone Exposure × Trial [habituated]	-0.028	0.044	289.999	-0.636	.525

^a Looking time ~ Age*Tone Exposure*Trial + (1 | Lab) + (1 | Subject)

TONE DISCRIMINATION IN INFANCY

List of Figures

Figure 1. Pitch contour and duration of the six Cantonese tones (right), including Chao numerals in the legends.

Figure 2. Summary of the literature investigating early native and non-native tone discrimination skills in monolingual and bilingual infants.

Figure 3. Looking time to habituated vs. novel test trials of the similar and distinct Cantonese tone contrasts by monolingual and bilingual infants (the data are collapsed across age groups, N = 416) (error bars represent 95% Confidence Intervals).

Figure 4. Looking time to habituated vs. novel test trials of the similar and distinct Cantonese tone contrasts by infants acquiring non-tone, pitch-accent, and tone languages (the data are collapsed across age groups, N = 416) (error bars represent 95% Confidence Intervals).

Figure 5. Looking time to habituated vs. novel test trials of the similar and distinct Cantonese tone contrasts by infants acquiring Cantonese vs. a tone language other than Cantonese (N = 155) (error bars represent 95% Confidence Intervals).

Figure 6. Looking time to habituated and novel trials as a function of bilingual infants' exposure to a tone language (0 to 75%) (the data are collapsed across age groups, N = 98) (shading represents 95% Confidence Intervals).

TONE DISCRIMINATION IN INFANCY

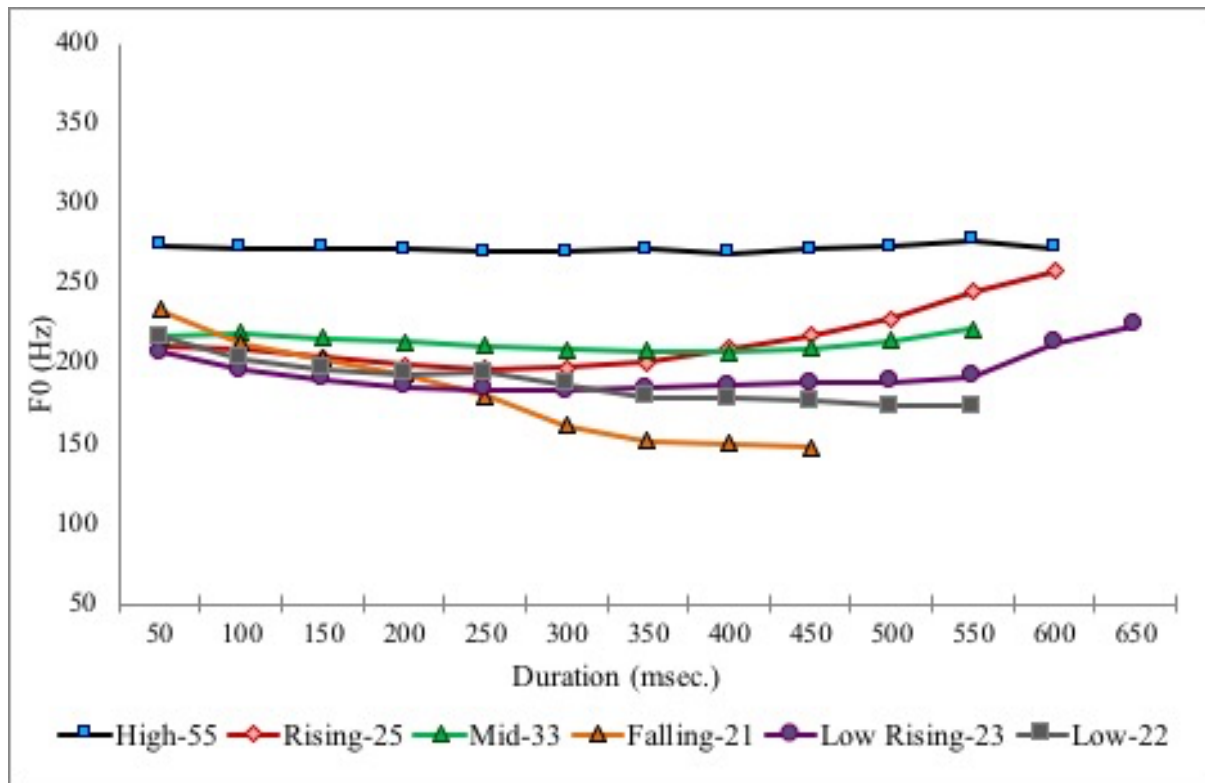


Figure 1. Pitch contour and duration of the six Cantonese tones (right), including Chao numerals in the legends.

STONE DISCRIMINATION IN INFANCY

	Maintenance	Facilitation	Decline	Non-discrimination	Decline and facilitation (U-shaped development)
Native (Monolingual)	<p>Mandarin or Cantonese: 6 and 9 months (Mattock & Burnham, 2006) [SALIENT 315-241 AND SUBTLE 315-21]</p> <p>Taiwanese Mandarin: 7 and 11 months (Tsao, 2017) [SUBTLE 35-214 & 35-51]</p> <p>Cantonese and Mandarin: 4 and 9 months (Yeung et al. 2013) [SUBTLE 25-33]</p>	<p>Taiwanese Mandarin: 7 and 11 months (Tsao, 2017) [SALIENT 55-214]</p> <p>Chinese: 6, 8, and 12 months (Singh et al. 2018; Singh et al. 2016) [SUBTLE 35-214 AND SALIENT 55-214]</p>			
Non-native (Monolingual)	<p>French: 4, 8 and 11 months (Shi et al. 2017) [SALIENT 55-51]</p> <p>Dutch: 5-17 months (Liu & Kager, 2014) [SALIENT: 55-21]</p> <p>Mandarin: 6 and 10 months (Cabrera et al., 2015) [SUBTLE 315-21]</p> <p>Dutch: 6, 9 and 12 months (Ramachers et al. 2018)</p>	<p>English: 7 and 11 months (Tsao, 2017) [SALIENT 55-214]</p> <p>Dutch: 4, 6 and 12 months (Chen & Kager, 2015) [SUBTLE 35-214]</p> <p>Dutch: 4 and 12 months (Chen, Stevens & Kager, 2017) [SUBTLE 35-214]</p> <p>English: 6, 8, and 12 months (Singh et al. 2018) [SUBTLE 35-214 AND SALIENT 55-214]</p>	<p>English and French: 4, 6, and 9 months (Mattock et al. 2008) [SUBTLE 315-21]</p> <p>English: 6 and 9 months (Mattock and Burnham, 2006) [SALIENT 315-241 AND SUBTLE 315-21]</p> <p>French: 6 and 10 months (Cabrera et al., 2015) [SUBTLE 315-21]</p> <p>French: 4, 8 and 11 months (Shi et al. 2017) [SUBTLE 315-214]</p> <p>English: 4 and 9 months (Yeung et al. 2013) [SUBTLE 25-33]</p>		<p>Dutch: 5 to 17 months (Liu & Kager, 2014) [SUBTLE 55-21(<i>compressed</i>)]</p> <p>German: 6, 9 and 18 months (Götz et al. 2018) [SALIENT 25-33]</p>
Native (Bilingual)	<p>Limburgian mixed: 6, 9 and 12 months (Ramachers et al. 2018).</p>			<p>Chinese-English: 6, 8, 12 months (Singh et al. 2018) [SUBTLE 35-214 AND SALIENT 55-214]</p>	
Non-native (Bilingual)	<p>Dutch mixed: 5 to 17 months (Liu & Kager, 2017) [SALIENT 55-51]</p>				<p>Dutch mixed: 5 to 17 months (Liu & Kager, 2017) [SUBTLE 55-21(<i>compressed</i>)]</p>

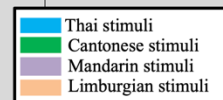


Figure 2. Summary of the literature investigating early native and non-native tone discrimination skills in monolingual and bilingual infants.

TONE DISCRIMINATION IN INFANCY

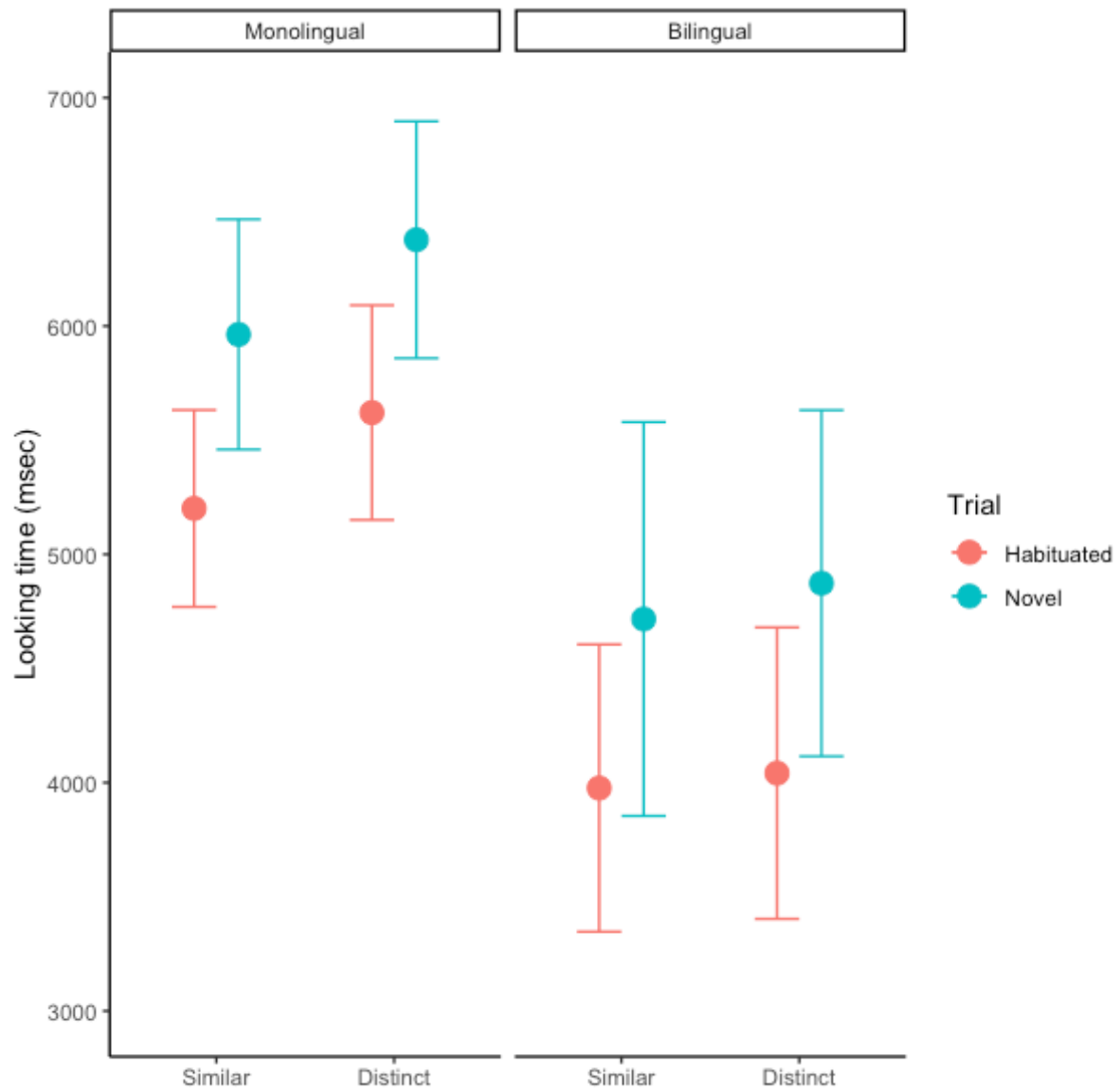


Figure 3. Looking time to habituated vs. novel test trials of the similar and distinct Cantonese tone contrasts by monolingual and bilingual infants (the data are collapsed across age groups, N = 416) (error bars represent 95% Confidence Intervals).

TONE DISCRIMINATION IN INFANCY

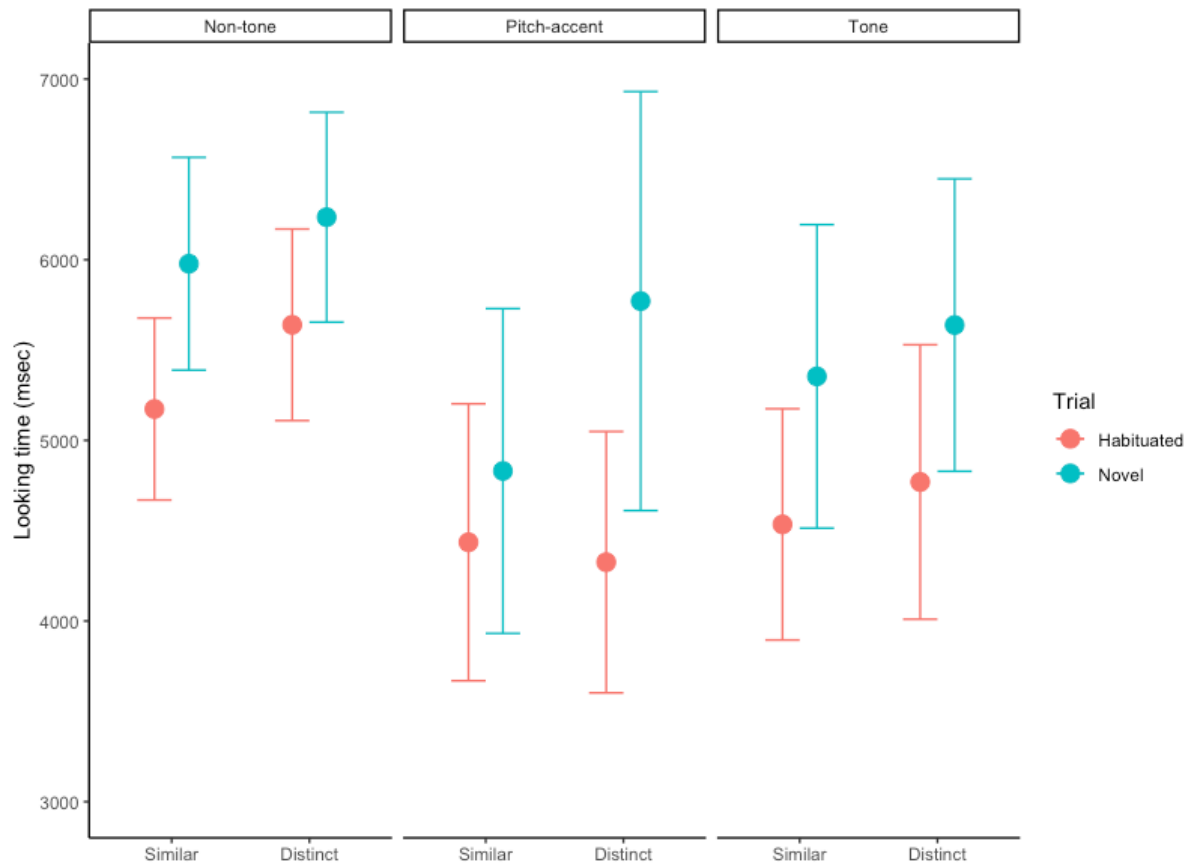


Figure 4. Looking time to habituated vs. novel test trials of the similar and distinct Cantonese tone contrasts by infants acquiring non-tone, pitch-accent, and tone languages (the data are collapsed across age groups, N = 416) (error bars represent 95% Confidence Intervals).

TONE DISCRIMINATION IN INFANCY

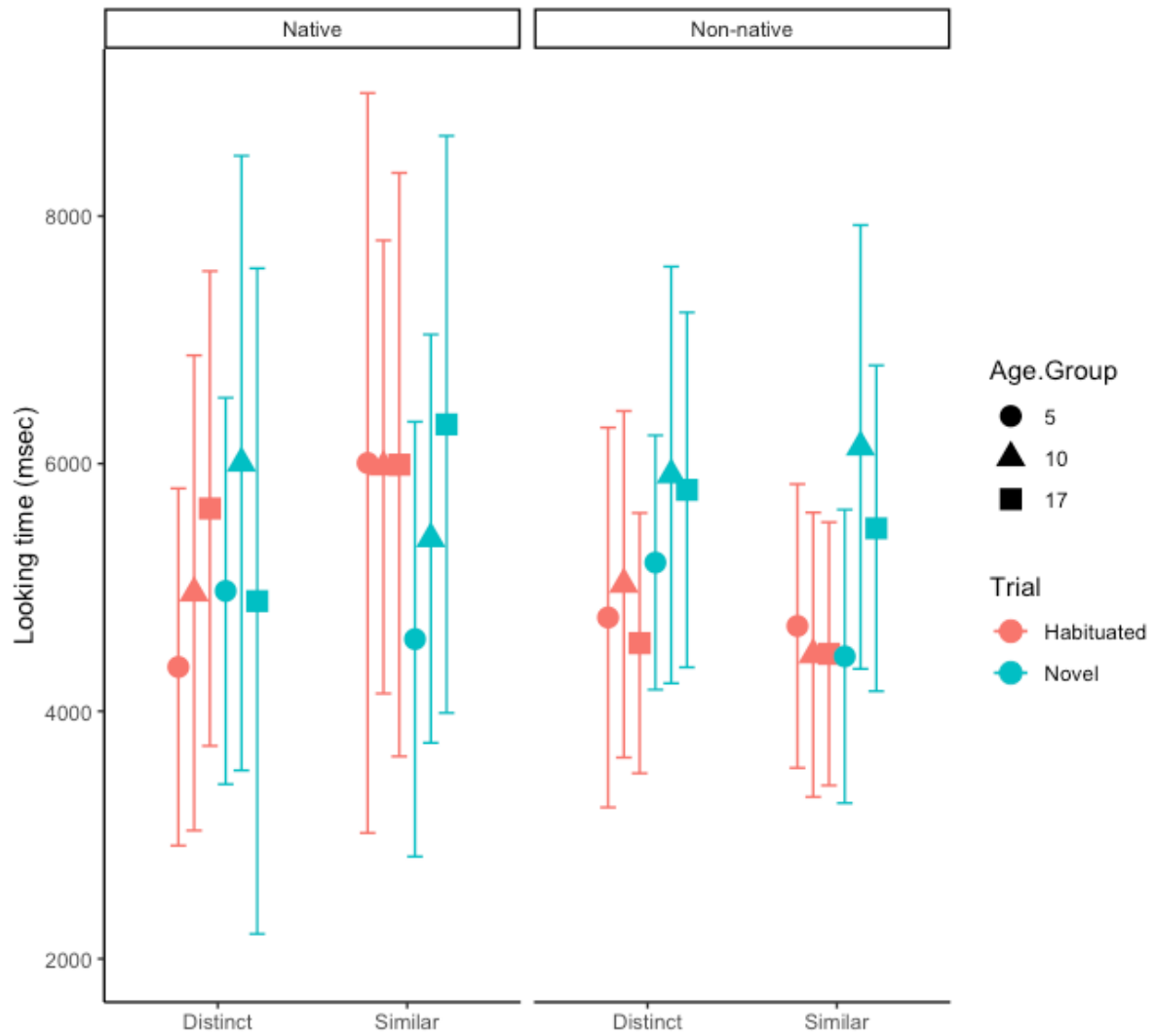


Figure 5. Looking time to habituated vs. novel test trials of the similar and distinct Cantonese tone contrasts by infants acquiring Cantonese vs. a tone language other than Cantonese (N = 155) (error bars represent 95% Confidence Intervals).

TONE DISCRIMINATION IN INFANCY

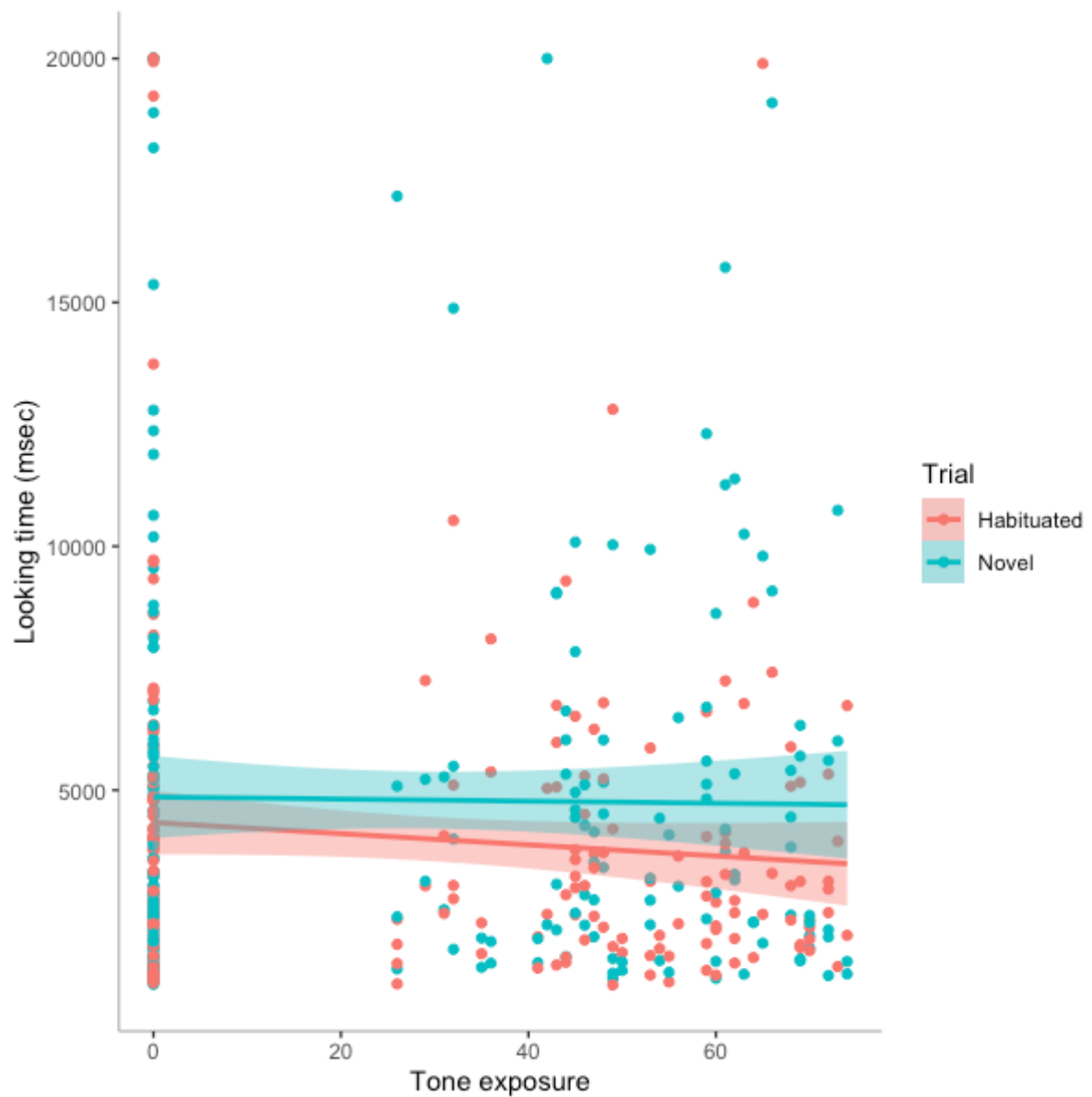


Figure 6. Looking time to habituated and novel trials as a function of bilingual infants' exposure to a tone language (0 to 75%) (the data are collapsed across age groups, N = 98) (shading represents 95% Confidence Intervals).