Journal of Experimental Child Psychology 164 (2017) 192–208

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Statistical learning of speech sounds is most robust during the period of perceptual attunement

Liquan Liu [a,b,*], René Kager [a]

[a] *School of Social Sciences and Psychology, Western Sydney University, Penrith, NSW 2751, Australia*
[b] *Utrecht Institute of Linguistics OTS, Utrecht University, 3512 JK Utrecht, The Netherlands*

ARTICLE INFO

Article history:
Received 19 May 2016
Revised 23 May 2017
Available online 4 July 2017

Keywords:
Infant
Language development
Distributional learning
Perceptual attunement
Speech perception
Lexical tone

ABSTRACT

Although statistical learning has been shown to be a domain-general mechanism, its constraints, such as its interactions with perceptual development, are less well understood and discussed. This study is among the first to investigate the distributional learning of lexical pitch in non-tone-language-learning infants, exploring its interaction with language-specific perceptual attunement during the first 2 years after birth. A total of 88 normally developing Dutch infants of 5, 11, and 14 months were tested via a distributional learning paradigm and were familiarized on a unimodal or bimodal distribution of high-level versus high-falling tones in Mandarin Chinese. After familiarization, they were tested on a tonal contrast that shared equal distributional information in either modality. At 5 months, infants in both conditions discriminated the contrast, whereas 11-month-olds showed discrimination only in the bimodal condition. By 14 months, infants failed to discriminate the contrast in either condition. Results indicate interplay between infants' long-term linguistic experience throughout development and short-term distributional learning during the experiment, and they suggest that the influence of tonal distributional learning varies along the perceptual attunement trajectory, such that opportunities for distributional learning effects appear to be constrained in the beginning and at the end of perceptual attunement. The current study contributes to previous research by demonstrating an effect of age on learning from distributional cues.

---

* Corresponding author at: School of Social Sciences and Psychology, Western Sydney University, Postbox 1797, Penrith, NSW 2751, Australia.

E-mail address: l.liu@westernsydney.edu.au (L. Liu).

http://dx.doi.org/10.1016/j.jecp.2017.05.013
0022-0965/© 2017 Elsevier Inc. All rights reserved.

Introduction

Statistical learning is the ability to extract statistical regularities and learn from the ambient environment (Kirkham, Slemmer, & Johnson, 2002; Saffran, Johnson, Aslin, & Newport, 1999). This domain-general learning mechanism plays a key role in various aspects of cognitive development across the lifespan (Conway & Christiansen, 2005; Escudero & Williams, 2014; Krogh, Vlach, & Johnson, 2013; Saffran, Aslin, & Newport, 1996; Turk-Browne, Jungé, & Scholl, 2005). In the linguistic domain, statistical learning has been shown to facilitate the acquisition of phonetic categories (Wanrooij, Boersma, & Van Zuijen, 2014) and words (Saffran et al., 1996). Sounds from different categories within a language overlap not only in speech production but also in the mental representation of phonetic space. To acquire native categories, infants must pay attention to and use the distributional frequency of speech sounds (Saffran et al., 1999). The current study focused on a specific type of statistical learning—learning from frequency distributions, arguably used in phonetic category formation among infants.

Since the introduction of the frequency distribution model by Maye, Werker, and Gerken (2002), several variations have been used to test how input statistics can alter rapid phonetic category learning and discrimination in infants. In their pioneering work, Maye and colleagues created a continuum of speech sounds based on a voiced versus voiceless unaspirated stop consonant contrast ([ta]–[da]) and exposed infants to the full continuum of stimuli arranged with unimodal or bimodal frequency distributions. Phonetic distributional learning studies typically involve these two statistical distributions, differing in the number of Gaussian peaks (tokens with relatively high frequency) along a phonetic continuum (Fig. 1). A unimodal distribution is characterized by one peak, corresponding to single category learning. In contrast, a bimodal distribution is marked by two different peaks, corresponding to the learning of two categories. Infants are familiarized with one of the two types of frequency distributions, followed by a measurement of their discrimination ability of tokens presented with equal frequency in the two distributions.

A recent study demonstrates very young infants' rapid distributional learning ability. Dutch infants of 2–3 months were presented with either a unimodal or bimodal distribution based on the English /æ/–/ɛ/ vowel contrast—a non-native contrast in Dutch—during sleep. Following a 12-min exposure phase, larger mismatch responses were found in the bimodally trained infants when compared with

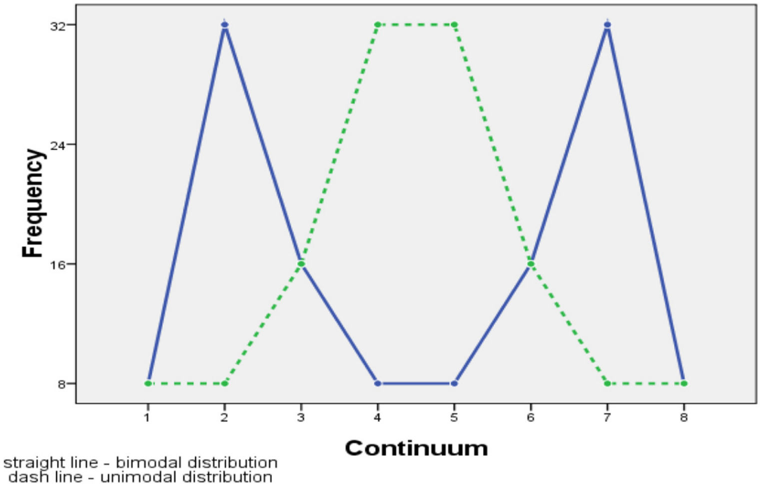

Fig. 1. Example of unimodal (dotted line) and bimodal (solid line) frequency distributions. In Maye et al. (2002), Token 1 represents the endpoint [da] stimulus and Token 8 represents the endpoint [ta] stimulus on the abscissa. The ordinate axis shows the presentation of frequency during familiarization.

those who received unimodal training, indicating better discrimination after bimodal distribution. In other words, a bimodal distribution facilitated young infants' perception of a non-native contrast (Wanrooij et al., 2014). Exposure to a unimodal distribution reduced 6- and 8-month-old English-learning infants' perceptual sensitivity to a non-native contrast (e.g., voiced vs. voiceless unaspirated alveolar stops [da]–[ta]) along the voice onset time continuum (Maye et al., 2002). Exposure to a bimodal distribution, on the other hand, allowed infants to extract abstract featural representations from the phonetic input and discriminate new non-native contrasts (e.g., prevoiced vs. unaspirated velar stops [ga]–[ka]) that were featural analogs of the trained contrast (Maye, Weiss, & Aslin, 2008).

Recent findings suggest that the amount of exposure to distributional cues required to produce learning effects may be gradient rather than absolute. At 10 months, English-learning infants did not demonstrate distributional learning of non-native speech sounds (e.g., voiced vs. voiceless unaspirated alveolar stops [da]–[ta]) when familiarized in the bimodal distribution. Nevertheless, indications of learning surfaced when the experimenters doubled the amount of exposure in the bimodal condition. In contrast, no learning effect was observed in the unimodal condition even when exposure in familiarization was doubled (Yoshida, Pons, Maye, & Werker, 2010).

At first glance, an increase in short-term distributional learning difficulty with age opposes expectations based on cognitive maturation (Elman, 1993; Gómez & Maye, 2005; Newport, 1988; Santelmann & Jusczyk, 1998), the construction and development of thought processes and cognitive functions. Specifically, older infants should have established better general cognitive abilities, such as selective attention and information processing, resulting in better performance than younger infants in a fast speech-learning task. Nevertheless, previous findings suggest otherwise. The issue arises as to the extent of age-related effects on short-term distributional learning ability and whether they are related to prior linguistic exposure. Compared with previous studies testing the same contrast (Maye et al., 2002, 2008), Yoshida et al. (2010) argued that rapid distributional learning of a non-native contrast may become more challenging for infants of an older age (e.g., at 10 months) as opposed to younger infants (e.g., at 6 months) due to several possibilities: (a) the greater amount of accumulated native input, not easily overcome by a short training period in the lab; (b) the development of native categories, leading to greater perceptual rigidity because the perception of non-native contrasts may require the learning of a second phonetic structure (Flege, 1995) and, hence, become more difficult (Best, 1995); and (c) a general age-related plasticity decline. All three explanations are relevant to perceptual attunement.

Perceptual attunement (also referred to as perceptual reorganization or narrowing) refers to the tuning-in process from universal sensitivity to the native inventory, typically during the first year of life (Anderson, Morgan, & White, 2003; Burns, Yoshida, Hill, & Werker, 2007; Flom, 2014; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker & Tees, 1984). It includes the maintenance and enhancement of the native phonetic categories as well as a decrease of sensitivity to non-native contrasts (Aslin & Pisoni, 1980) across ages. The mechanisms underlying perceptual attunement remain unclear. Statistical learning has been argued to be one of the main mechanisms that drive this developmental change (Maye et al., 2002). Furthermore, an additional maturational component, a critical or "optimal" period (Werker & Hensch, 2015; Werker & Tees, 2005), may contribute to this category formation process along the developmental trajectory. With accumulated long-term native language exposure, older infants establish more stabilized native phonetic representations and are prone to ignoring non-native phonetic distinctions more than younger infants (Werker & Tees, 1984) unless the distinctions are highly perceptually salient (Best, McRoberts, & Sithole, 1988). Although perceptual attunement for consonants occurs near the end of the first year after birth (Werker & Tees, 1984), infants demonstrate native-like perception for vowels (Kuhl et al., 2006; Polka & Werker, 1994) and lexical tones, the word-level pitch variations that distinguish meaning, by the age of 9 months (Burnham et al., 2014).

Lexical tones exist in tone languages (e.g., Mandarin Chinese) but are lacking in non-tone languages (e.g., Dutch). Different from consonants and vowels, tones contrast via a suprasegmental feature—pitch. Non-tone language-speaking adults perceive tonal contrasts acoustically rather than categorically, which is how they perceive consonant contrasts (Chen, Liu, & Kager, 2015, 2016; Hallé, Chang, & Best, 2004). On the one hand, tone language learners maintain their sensitivity to native and even non-native lexical tonal contrasts, as found in Cantonese infants of 4 months, Yorùbá infants of 6 months, and Mandarin Chinese infants of 4–9 months (Harrison, 2000; Yeung, Chen, & Werker, 2013). On the other hand, non-tone-language-learning infants demonstrate initial sensitivity to lexical tones at birth (Nazzi, Floccia, & Bertoncini, 1998) but undergo a decline of sensitivity during the first year of life. Reduced sensitivity to lexical tones in Thai, Cantonese, and Mandarin Chinese has been found in English- and Dutch-learning infants between 6 and 9 months of age (Mattock, Molnar, Polka, & Burnham, 2008) regardless of the fact that these infants have experience with intonation–pitch modulations at a sentence level. Liu and Kager (2014) tested Dutch infants' ability to discriminate two Mandarin lexical tonal contrasts via a habituation paradigm. Results showed that infants discriminated the non-native contrasts at 5 or 6 months and that the deterioration of their tonal sensitivity varied as a function of acoustic salience. At 11 and 14 months, infants did not show discrimination of a manipulated tonal contrast with reduced salience. When tested on a distributional learning task, 11-month-old infants exposed to a bimodal distribution demonstrated increased discrimination ability to non-native tones.

The current experiment instantiated a learning scenario of non-native tones because non-tone-language-learning infants do not have previous experience with non-native tones, nor do they have tonal categories in place (Liu, L., & Kager, R. (2011); Liu, L., & Kager, R. W. J. (2011); Liu & Kager, 2013, 2017). Although non-tone-language-learning infants' sensitivity to lexical tones deteriorates during the first year after birth, it remains unclear whether certain fast learning conditions may enhance (e.g., a bimodal frequency of tonal distribution; Liu & Kager, 2014) or impede (e.g., unimodal; Maye et al., 2002) infants' perception of a non-native contrast differing in suprasegmental features. Crucially, the interaction between such fast learning and the long-term non-tonal experience (e.g., perceptual attunement) has not yet been revealed along the developmental trajectory.

To advance our knowledge on the effect of fast learning of a non-native contrast on perceptual sensitivity under various perceptual attunement stages, infant distributional learning of non-native lexical tones across ages was examined. Based on the outcomes of previous literature on tonal perceptual attunement (e.g., Mattock & Burnham, 2006), three age groups were selected for investigation: one in the beginning of the attunement period (at 5 months), one at the end of the attunement period (at 11 months), and one in the second year after birth (14 months) when infants have established more robust native phonetic categories (Dietrich, Swingley, & Werker, 2007).

We predicted that infants' non-native tone discrimination would relate to their exposure to and learning from the distributional information with which they were familiarized. More specifically, we predicted that infants in the bimodal condition would show a greater facilitation effect when discriminating the non-native tone contrast than infants in the unimodal condition. Provided that statistical learning is a domain-general mechanism available across the life span, the effects of bimodal facilitation and/or unimodal inhibition should surface at all ages. Nevertheless, empirical findings appear to suggest that infant distributional learning ability declines during the first year of life (Yoshida et al., 2010). In addition, perceptual attunement has been shown to moderate learning. Infants' reduced sensitivity under the late perceptual attunement period may negatively affect their learning ability of non-native contrasts from a bimodal distribution. This would predict a weakened bimodal facilitation effect at a later stage (11 and/or 14 months) along the perceptual development. On the other hand, infants in the early phase of perceptual attunement (5 months) would keep their tonal sensitivity, although exposure to a unimodal condition may lead to sensitivity impediment.

Method

Participants

A total of 134 typically developing 5-, 11-, and 14-month-old monolingual Dutch infants from comparable socioeconomic backgrounds participated in the study. All infants were recruited through the Babylab database of Utrecht Institute of Linguistics OTS, Utrecht University. Data of 88 infants (46 male and 42 female) were incorporated into the analysis. The remaining 46 infants were excluded from the analysis due to crying or excessive fussiness ($n = 14$), parental interference ($n = 1$), looking time not reaching 60% of the total looking time during the familiarization phase[1] ($n = 11$), failure to habituate (after 25 trials) during the habituation phase ($n = 12$), looking time less than 2 s in both dishabituation trials ($n = 4$), and outliers (looking time in dishabituation differing more than 2 standard deviations from the mean of the familiarization condition group) ($n = 4$). The detailed dropout numbers across age groups and conditions are listed in Appendix 'Dropout numbers across ages and conditions'. A dropout rate of 34% indicates that the task may be cognitively demanding compared with other studies. In each age group, infants were randomly assigned to a unimodal or bimodal condition. Table 1 illustrates the detailed demographic information per age group and modal condition.

Stimuli

Multiple natural tokens of the Mandarin high-level (T1) versus high-falling (T4) tone bearing syllables /ta/ were produced by a female Mandarin speaker in a soundproof booth at the phonetics lab of Utrecht Institute of Linguistics OTS, Utrecht University. Tokens were recorded using the computer program Audacity[2] via a microphone (active speaker Genelec 1029A, sampling rate at 44,100 Hz). Four tokens of /ta/ in both tones were selected to increase within-speaker variation and facilitate tonal learning (Chen & Kager, 2015), thereby forming four pairs of level-falling tones. All selected tokens were intensity and duration controlled via the computer program PRAAT (Boersma & Weenink, 2009). Each pair (/taT1/–/taT4/) was further divided into six equidistant in-between tokens, creating an eight-step continuum (Fig. 2). This was done in PRAAT through the following steps. First, four points in time (0%, 33%, 67%, and 100%) were marked on the pitch contours of each pair. Second, the distances (in Hz) between the correspondent points of each pair were divided into seven equal spaces, creating 24 new in-between points on six layers. Third, new pitch tokens were generated via linking the four correspondent in-between points on the same layer. This forms a continuum of eight steps, including the endpoint contours from /taT1/ (Step 1) to /taT4/ (Step 8). A total of 32 stimuli were generated for the four continua as multiple individual tokens (Appendix 'F0 values at four points of the stimuli along /ta1/–/ta4/ continua'). The T1–T4 contrast in Mandarin Chinese has been shown to be discriminated by tone-language (Huang & Johnson, 2010) and non-tone-language (Liu, Chen, & Kager, in press) adult listeners and, hence, was considered as suitable for testing infants learning a non-tone language in the current study.

Design

Infants were randomly assigned to a unimodal or bimodal condition and listened to a sequence of tones that matched the condition. The two conditions differed only in the frequency distribution (one vs. two Gaussian peaks) along the phonetic continuum but not the total amount of tonal input (128 tokens) or duration (180 s). Specifically, stimuli near the peripheral sides of the continuum were presented with the highest frequency in the bimodal condition, whereas those from the central positions were presented most frequently in the unimodal condition. After the familiarization of the distribution condition, all infants were tested on their discrimination of tokens from Steps 3 and 6. These two steps overlapped in frequency of occurrence between unimodal and bimodal familiarization conditions; thus, any potential differences between the two familiarization conditions should not be due to the differences in exposure to the specific test stimuli. Infants were habituated on tokens of one step and were tested on those of the other step.

Procedure

During the experiment, infants sat on their caretaker's lap in the testing booth, facing the screen and the camera. No visual or auditory interference was observable in the booth other than the stimuli used in the test. The test was conducted using a computer program (ZEP; Veenker, 2007). An

---

[1] Recent infant studies require infants to accrue a certain amount of looking time during the familiarization phase (e.g., Yeung & Werker, 2009) or a minimum percentage of looking to have their data included in the analysis. The current percentage conforms to previous studies (Liu & Kager, 2014, 2017), allowing for a direct comparison of outcomes across studies.

[2] Audacity open resource: http://audacity.sourceforge.net.

Table 1
Participant information per age group and modal of distribution.

Age	Mean (*SD*) age in days		Number of boys/girls	
	Unimodal	Bimodal	Unimodal	Bimodal
5 months	169 (12)	173 (16)	5/10	3/11
11 months	355 (8)	354 (12)	9/5	8/7
14 months	421 (21)	441 (18)	11/5	10/4

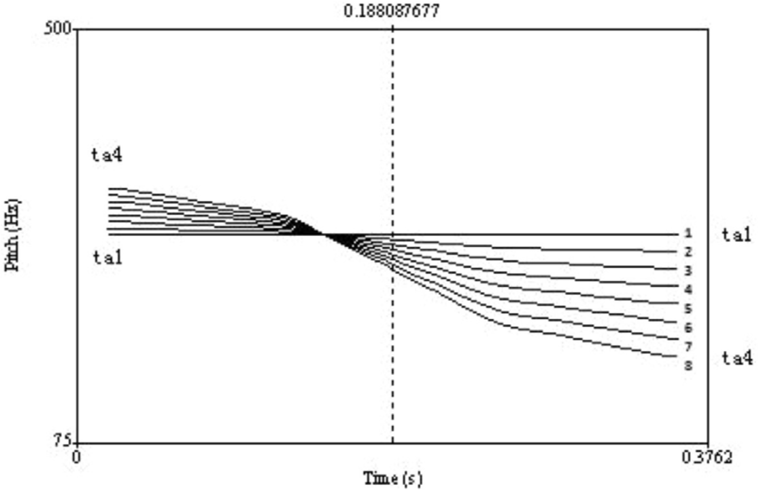

Fig. 2. Eight pitch contours along a /ta1/–/ta4/ continuum.

experimenter, seated in the testing room adjacent to the testing booth, observed the experiment through a closed-circuit TV and used a button box to record infants' looking patterns. Both experimenters and caretakers heard masking music through headphones and could not hear the stimuli presented to infants.

A similar distributional learning paradigm as used in Maye et al. (2008)[3] was adopted. The paradigm consisted of a familiarization phase, followed by a habituation procedure that consisted of two phases: habituation and test. The dependent variable was infant looking time. During familiarization, infants listened to a sequence of tones in the unimodal or bimodal distribution, and toy pictures were displayed on the screen (Appendix 'Related figures in procedure'). The acoustic tokens within each distributional condition were randomized. The total familiarization time was 180 s.

Following familiarization, infants went through an infant gaze-controlled habituation procedure and were tested on their discrimination of a contrast previously heard during familiarization. Within each group, half of the infants were habituated on tokens of Step 3 and tested on tokens of Step 6, and the other half had the habituation and testing tokens reversed. The effect of habituation order did not yield a significant difference; hence, all data were aggregated for analyses. The habituation criterion was reached when the sum of looking time of the last three habituation trials was lower than 65% of that of the first three habituation trials. Tokens of Steps 3 and 6 had the same frequency of occurrence in both unimodal and bimodal conditions. The design eliminated the possibility of attributing

---

[3] Some settings were adjusted in the current study compared with Maye et al. (2008). The visual stimuli during familiarization were images instead of silent videos. The habituation criterion was 65% in the current study instead of 50%. These criteria conform to our previous studies testing infants' pure discrimination of the contrast (Liu & Kager, 2014, 2017), allowing for a direct comparison of outcomes across studies.

infants' potential looking time differences during the habituation phase to unequal frequency of distribution during familiarization (Maye et al., 2008). The looking time recovery during the phase change (from the final two habituation trials to the two test trials) indicated discrimination.

Trials were discontinued when infants looked away from the monitor for more than 2 s or when the maximum trial length of 25 s was reached. A static "bull's eye" was presented on the screen in each habituation trial (Appendix 'Related figures in procedure'). The interstimulus interval was set as 1 s throughout each phase. The experiment ended with a Dutch child song accompanied with toy pictures to increase the friendliness of the experiment.

Results

Habituation phase

Infants' means and standard deviations of looking time differences between the first three and last three habituation trials are listed in Table 2. Infants' mean looking times between the first three and last three habituation trials were analyzed to assess habituation. A repeated-measures analysis of variance (ANOVA) was conducted with infants' mean looking times of the first three and last three habituation trials as the within-participants variable and age (5, 11, or 14 months) and distribution condition (unimodal or bimodal) as the between-participants factors. The main effect of looking time, $F(1, 82) = 139.167$, $p < .001$, $\eta^2 = .629$, was significant. The effect of age on looking time, $F(2, 82) = 1.388$, $p = .255$, $\eta^2 = .033$, the effect of distribution condition on looking time, $F(1, 82) = 0.030$, $p = .863$, $\eta^2 < .001$, and the interaction among age, distribution condition, and looking time, $F(2, 82) = 0.220$, $p = .803$, $\eta^2 = .005$, did not approach significance. These results indicated that infants across ages and familiarization conditions were habituated to the contrast.

Test phase

Infants' mean looking times between the last two habituation trials and the two test trials were analyzed to assess discrimination. Among all participants, 4 infants' mean looking times in test trials differed by more than 2 standard deviations from the mean of their familiarization condition and were considered as outliers and excluded from the analysis.

A repeated-measures ANOVA was conducted with infants' mean looking times as the dependent variable, phase of the experiment (the last two habituation trials or the two test trials) as the within-participants factor,[4] and age (at 5, 11, and 14 months) and distribution condition (unimodal and bimodal) as the between-participants factors. The main effect of looking time, $F(1, 82) = 16.739$, $p < .001$, $\eta^2 = .170$, and the effect of age on looking time, $F(2, 82) = 4.277$, $p = .017$, $\eta^2 = .094$, were significant. The effect of distribution condition on looking time, $F(1, 82) = 0.470$, $p = .495$, $\eta^2 = .006$, and the interaction among age, distribution condition, and looking time, $F(2, 82) = 1.762$, $p = .178$, $\eta^2 = .041$, did not approach significance. These results indicate that age was a significant factor in the analysis and that infants from various age groups performed differently (Fig. 3). Pairwise comparisons revealed differences among age groups; the 5-month-olds performed significantly different from the 11-month-olds ($p = .005$) and the 14-month-olds ($p = .006$), whereas the latter two age groups did not differ ($p = .895$).

To further examine the effect of distributional learning across ages, similar analyses were conducted within each age group. At the age of 5 months, there was a significant main effect of looking time, $F(1, 27) = 11.507$, $p = .002$, $\eta^2 = .299$. Nevertheless, the effect of distribution condition on looking time was not significant, $F(1, 27) = 0.119$, $p = .732$, $\eta^2 = .004$. Infants under both conditions successfully discriminated the tonal contrast. Splitting data by distribution condition, paired-samples t tests (two-tailed hereafter) with looking time as the dependent variable showed that infants under both

---

[4] Because no representation of habituation tones appeared in test trials, the results could be due to regression to the mean following attainment of the habituation criterion. However, this interpretation is unlikely given infants' performances across ages.

Table 2
Mean (*SD*) looking time differences between the last three habituation trials and the first three habituation trials across ages and conditions.

Age	Mean (*SD*) looking time difference	
	Unimodal	Bimodal
5 months	9.83 (9.62)	8.79 (6.77)
11 months	6.48 (2.53)	7.04 (5.84)
14 months	8.66 (6.46)	9.89 (6.86)

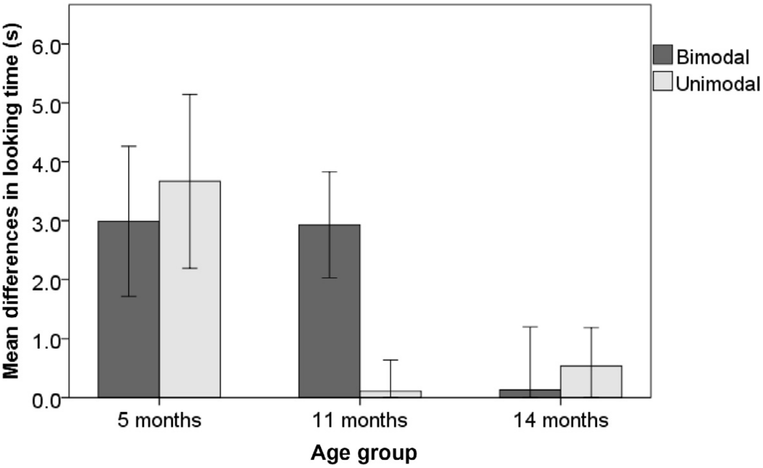

Fig. 3. Mean looking time differences between the end of habituation and the test trials for infants in the bimodal and unimodal conditions. The differences (in seconds) indicate attention recovery during the phase change. Error bars = ±1 standard error.

the unimodal condition, $t(14) = -2.483$, $p = .026$, and the bimodal condition, $t(13) = -2.349$, $p = .035$, presented significant looking time increases from the end of habituation trials to the test trials.

At 11 months, both the main effect of looking time, $F(1, 27) = 8.182$, $p = .008$, $\eta^2 = .233$, and the effect of distribution condition on looking time, $F(1, 27) = 7.054$, $p = .013$, $\eta^2 = .207$, were significant, indicating differences in infant looking patterns across conditions. Splitting the data between the two distribution conditions, only infants in the bimodal condition, $t(14) = -3.255$, $p = .006$, but not the unimodal condition, $t(13) = -0.206$, $p = .840$, discriminated the contrast during the habituation phase.

At 14 months, neither the main effect of looking time, $F(1, 28) = 0.299$, $p = .589$, $\eta^2 = .011$, nor the effect of distribution condition on looking time, $F(1, 28) = 0.112$, $p = .740$, $\eta^2 = .004$, was significant. Infants under neither condition discriminated the contrast at this age (Fig. 3). This was true for infants under the unimodal condition, $t(15) = -0.824$, $p = .423$, and the bimodal condition, $t(13) = -0.120$, $p = .906$.

Discussion

This study tested Dutch infants of 5, 11, and 14 months on their distributional learning ability for a non-native tonal contrast. Results unveiled that this statistical learning mechanism was most effective at 11 months, when infants in the bimodal condition, but not the unimodal condition, demonstrated discrimination. Younger infants at 5 months discriminated the contrast regardless of whether they were familiarized to a unimodal or bimodal distribution. Older infants at 14 months did not show

Table 3
General discrimination ability of the same tonal contrasts by Dutch infants under different ages and conditions.

Condition	Age		
	5 months	11 months	14 months
Bimodal distribution	Yes	Yes	No
Unimodal distribution	Yes	No	No
Pure discrimination	Yes	No	No

Note. Infants discriminated the non-native contrast across conditions at 5 months. At 11 months, a bimodal distribution facilitated discrimination. At 14 months, infants no longer discriminated the contrast in any condition.

discrimination regardless of the distributional condition with which they were familiarized. These results indicate differential efficacy of statistical learning at different points in development. The increasing difficulty of distributional learning with age was further discussed in light of previous data on infants' pure discrimination of the same contrast (Liu & Kager, 2014). Results showed that only infants of 5 months, but not the other two age groups, discriminated the non-native tonal contrast, conforming to findings on perceptual attunement in non-tone-learning infants from previous studies. The results from distributional learning and discrimination are listed in Table 3.

As Table 3 illustrates, infants present general sensitivity to non-native lexical tones at the early stage of tonal perceptual attunement. Distributional learning at 5 months appeared to have a limited impact on infants' general perception of tones. Crucially, statistical exposure to a unimodal distribution did not hinder infant sensitivity at 5 months. We hypothesize that infants discriminate the acoustic features of the contrast, and their robust initial tonal sensitivity may outweigh input statistics at this stage. Because acoustic salience plays an important role in early speech perception and perceptual attunement (Kuhl et al., 2008), a tonal contrast with less prominence or salience is predicted to lead to an inhibition effect on infants' perceptibility in the unimodal condition.

Non-tone-language-learning infants pay little attention to the phonetic details of non-native tones after 9 months (Mattock & Burnham, 2006). Likewise, 11-month-old Dutch infants failed to discriminate the contrast without being preexposed to any tonal distributions (Table 3). Regardless of this sensitivity decline, the current experiment illustrates that infants' discrimination of non-native tones varied as a function of distributional information. Exposure to the unimodal condition did not facilitate the discrimination of the contrast, but exposure to the bimodal condition did. The overall findings suggest that it is the bimodal exposure that enhances perception. The asymmetrical perceptual patterns between infants from different modal conditions resemble earlier studies showing that infants' discrimination of non-native contrasts can be pushed "up" only when given the "right" type of (bimodal) exposure (Maye et al., 2002, 2008; Wanrooij et al., 2014; Yoshida et al., 2010).

At 14 months, Dutch infants failed to discriminate the non-native tonal contrast at each of the three conditions. These infants appear to have established a relatively reliable native sound inventory. The cumulative effects of continuous exposure to a native language that lacks tonal categories during the first 14 months of life may no longer be overcome by 3-min fast learning of tonal distributions. The increasing distributional learning difficulty, reflected in infants' lack of sensitivity even habituated in the bimodal condition, is also observed in 10-month-olds with a different contrast (Yoshida et al., 2010). The different time windows across the two studies may reflect the relative learning difficulty due to the prominence of contrast acoustic properties. We hypothesize that pitch contrasts may be acoustically more salient than consonant contrasts tested in previous studies (Maye et al., 2002., 2008; Yoshida et al., 2010), resulting in potential better perceptibility and statistical learning outcomes at a stage following perceptual attunement for tone. Future studies should examine this issue by varying the amount of exposure to distributions at different ages.

The outcomes of the current study provide cross-sectional data illustrating the dynamic interaction between long-term perceptual attunement and short-term distributional learning of non-native speech contrasts. Age matters in the effectiveness of distributional frequency information for the discrimination of non-native contrasts during infancy. We hypothesize that perceptual attunement, the general age-related plasticity decline, affects rapid distributional learning. The size of the learning effects may be subject to how relevant the target is to infants' previous knowledge established from the environment. During the course of language acquisition, learners may gradually become less sensitive to distributional statistics of non-phonemic variation or linguistic properties irrelevant to the native language given their accumulated experience. Non-tone-language-learning infants may spend fewer cognitive resources, such as higher level attentional processes, on computing linguistic tonal variations because it is irrelevant for learning in their environment. This strategy would facilitate infants' acquisition of native phonemic inventories. Alternatively, older infants' fast learning ability may decline as a function of the decreasing perceptibility of a non-native contrast. Different from the first hypothesis, this explanation taps into the general decrease of perceptual sensitivity to irrelevant or infrequent information. Acoustically contrastive elements that matter little to native language development will be less well perceived (e.g., categorical perception).

A third explanation for older infants' overall lack of sensitivity across conditions (Table 3) is that phonetic learning may be associated with cues beyond statistical learning (Potter, Wang, & Saffran, 2017; Singh, Nestor, Parikh, & Yull, 2009; Yeung, Chen, & Werker, 2014). At different ages, input distribution might not always be the most robust cue among other learning cues and strategies. Although statistical learning is considered as a general learning mechanism across ages and domains, this learning mechanism is not without constraints on the type of input and/or on learners (Krogh et al., 2013). Infants' reliance on cues changes during the course of development (Lany & Saffran, 2011; Shukla, Nespor, & Mehler, 2007). In the speech domain, compared with 5- and 7-month-olds, 9-month-olds rely more on input phonology than on syllable co-occurrence to segment words when both cues are provided (Thiessen & Erickson, 2012; Thiessen & Saffran, 2003). Likewise, 6-month-olds have been shown to concentrate more on prosodic boundaries than on syllable co-occurrence (Shukla, White, & Aslin, 2011). The shift of infants' cue weighting may reflect their online processing preferences. In other words, the weighting of specific cues may increase at a specific developmental period if certain cues facilitate learning better than the others. Along similar lines, factors affecting the exposure to auditory and visual information, such as (joint) attention (Conboy, Brooks, Taylor, Meltzoff, & Kuhl, 2008; Ong, Burnham, & Escudero, 2015), facial articulation (Teinonen, Aslin, Alku, & Csibra, 2008), object labeling (Yeung & Werker, 2009), and live social interaction (Kuhl, Tsao, & Liu, 2003) may also influence rapid distributional learning. In a study in which infants learned from exposure to a new language where various cues were presented (Kuhl et al., 2003), distributional frequency alone was insufficient for learning, whereas social interaction produced robust learning of non-native contrasts. During the second year after birth, the rate at which children acquire new words accelerates drastically. Infants may concentrate more on the use of lexicon and less on the incoming distributions (McMurray, 2007), in line with the claim that infants access different focus points at different developmental stages (Golinkoff, Can, Soderstrom, & Hirsh-Pasek, 2015; Hollich, Hirsh-Pasek, & Golinkoff, 2000). The cue reliance constraint explanation can be tested in future studies by adding additional learning cues to the distributional paradigm.

One point worth mentioning is the potential impact of native suprasegmental features on non-native tone perception. Do infants' native stress and intonation acquisition facilitate their perception of non-native tones? Stress is the relative emphasis that may be given to syllables or words. Stressed syllables and words may result in changes in pitch. Infants' preference for the predominant native stress patterns surfaces as early as 9 months after birth (Jusczyk, Cutler, & Redanz, 1993). Intonation is characterized to a large extent by pitch variation at an utterance level. A large number of studies report infants' sensitivity to prosodic cues at the utterance level during the first year after birth (Butler, Vigário, & Frota, 2016; Frota, Butler, & Vigário, 2014; Männel & Friederici, 2009; Nazzi, Nelson, Jusczyk, & Jusczyk, 2000; Pannekamp, Weber, & Friederici, 2006; Seidl & Cristià, 2008; Sundara, Molnar, & Frota, 2015), within which the socioemotional information is embedded (Flom & Bahrick, 2007; Grossmann, Striano, & Friederici, 2006). Nevertheless, Dutch infants' native trochaic stress pattern and their sensitivity to intonation over time do not seem to facilitate the perception of the tonal contrast across conditions. There is no evidence showing that older infants benefit from their richer knowledge of prosody and perform better than the younger ones.

We hypothesize that a potential first-language facilitation transfer effect from native intonation might still exist because many properties of intonation resemble those of tone. However, the effect does not surface due to the long acquisition trajectory of intonation. For example, one's native intonation system is not stabilized even after the second year due to its complex linguistic use. Infants' intonation production of language-specific pitch patterns does not become adult-like until 21 months (Frota & Vigário, 2008), and the pitch range in statement intonation was not well controlled even after 2 years of life (Vanrell, Prieto, Astruc, Payne, & Post, 2010). Acquisition of intonation is a cumulative process that involves integration of knowledge about pitch contours, grammatical structures, and pragmatic meanings. For instance, the acquisition of socioemotional knowledge embedded in prosody continues even at 9 years of age (Nelson & Russell, 2011). The native language transfer effect may emerge at a later stage, assisting non-native tone perception and subsequently learning (Singh & Chee, 2016). We leave this issue open for future research.

Infants' overall sensitivity to non-native tones implies the importance of long-term experience on fast learning in the lab setting. Although the statistical learning mechanism is age and domain general, it is by no means the only one in action. Our findings indicate that although statistical learning may very well be a learning mechanism that is available across the life span, it is not equally effective at all ages. Rather, infants may develop multiple learning strategies that fit the individual learning environments throughout development.

Conclusion

This study addressed the relationship between non-tone-language-learning infants' cross-age tone perception and distributional learning, testing infants' learning of non-native lexical tones under two types of frequency distributions. The unimodal distribution did not hinder 5-month-olds' non-native tonal discrimination at the early stage of perceptual attunement. At 11 months, the bimodal distribution, but not the unimodal distribution, appeared to facilitate discrimination, possibly when infants required extra support from the surroundings. At 14 months, the bimodal exposure was no longer sufficient to assist infants to achieve discrimination. The fact that prior to 11 months of age infants continued to discriminate the tones even following unimodal familiarization, and that bimodal familiarization could not enhance discrimination after this age, provides evidence not for the suitability of the procedure for the age of infants but rather for a differential impact of fast learning during various periods of development.

The current study contributes to the existing literature by demonstrating an effect of age on learning from distributional cues. Results suggest an interplay between infants' long-term linguistic experience throughout development and short-term 3-min distributional learning during the experiment, and they suggest that the opportunities for a distributional learning effect are constrained along infants' tonal developmental trajectory, revealing the role of input under perceptual attunement; the deterioration of non-native tonal sensitivity is a gradual and flexible process (Werker & Hensch, 2015; Werker & Tees, 2005), influenced by age, input frequency, and acoustic properties (Kuhl et al., 2003; Yoshida et al., 2010). Additional cues may apply for successful acquisition at different developmental stages (McMurray, Aslin, & Toscano, 2009). A future question is perhaps to explore ways in which to maintain the strength of distributional learning of speech sounds at a relatively later age, after the perceptual attunement period.

Acknowledgments

We dearly thank Ben Thompson, Brendan Devers, Illeana Grama, Jessica Bazouni, Jessie Nixon, Lawrence Lo, Mieke Du Toit, Muna Pohl, Murray Schellenberg, Nick Pijnappels, Paola Escudero, Rachel Gargiulo, Tianlin Wang, Tomas Lentz, and Wei Wang for commenting on the manuscript. We sincerely thank the Babylab group and experimental phonology group members of the Utrecht Institute of Linguistics OTS, Utrecht University, for their valuable feedback. Special thanks go to the reviewers of the *Journal of Experimental Child Psychology*. Last, but not least, we are in debt to all families that participated in our research in The Netherlands. A subset of the 11-month-old infants' data was used in Liu (2011), 2014).

Appendix. Dropout numbers across ages and conditions

Dropout reason[a]	Condition and age					
	Unimodal			Bimodal		
	5 months	11 months	14 months	5 months	11 months	14 months
1	3	0	6	2	0	3
2	0	0	0	0	1	0
3	0	2	1	3	4	1
4	2	3	0	4	1	2
5	2	1	0	0	1	0
6	1	0	0	1	1	1

[a] Dropout reasons: The numbers follow the exclusion criteria in the "Participants" section: 1 = crying or excessive fussiness (*n* = 14); 2 = parental interference (*n* = 1); 3 = looking time not reaching 60% of the total looking time during the familiarization phase (*n* = 11); 4 = failure to habituate (after 25 trials) during the habituation phase (*n* = 12); 5 = looking time less than 2 s in both dishabituation trials (*n* = 4); 6 = outliers (looking time in dishabituation differing more than 2 standard deviations from the mean of the familiarization condition group) (*n* = 4).

Appendix. F0 values at four points of the stimuli along /ta1/–/ta4/ continua

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

Continuum A

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

Continuum B

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

Continuum C

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

Continuum D

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

Appendix. Related figures in procedure

See Figs. A1–A4.

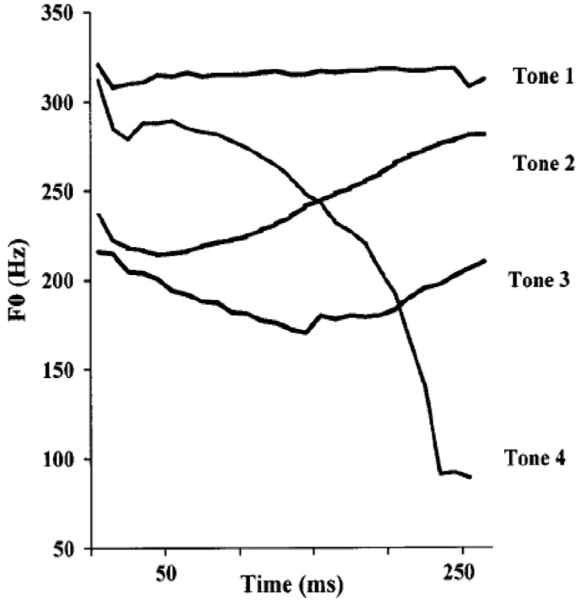

Fig. A1. Tones in Mandarin Chinese. *Source:* Wang, Jongman, and Sereno, 2001.

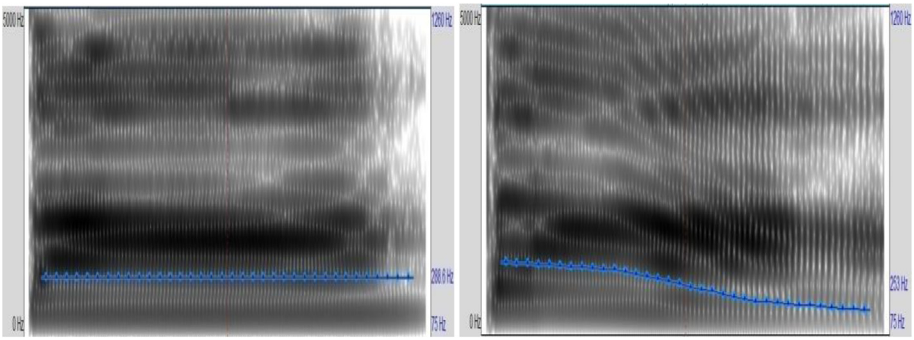

Fig. A2. Spectrograms of the tonal contrast (left: /ta/ T1; right: /ta/ T4).

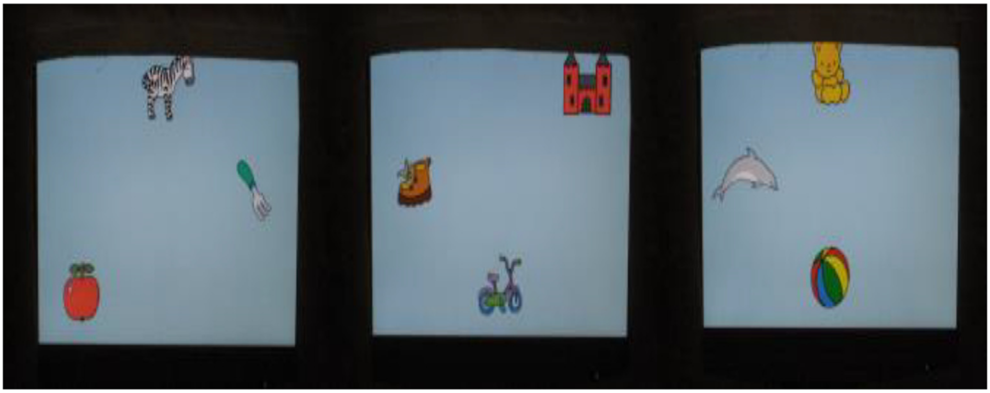

Fig. A3. Visual stimuli during the familiarization phase.

Fig. A4. Visual stimulus during the habituation and test phases.

References

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

Aslin, R. N., & Pisoni, D. B. (1980). Some developmental processes in speech perception. *Child Phonology, 2*, 67–96.

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

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Timonium, MD: York Press.

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

Burnham, D., Kasisopa, B., Reid, A., Luksaneeyanawin, S., Lacerda, F., Attina, V., ... Webster, D. (2014). Universality and language-specific experience in the perception of lexical tone and pitch. *Applied Psycholinguistics, 77*, 571–591.

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

Butler, J., Vigário, M., & Frota, S. (2016). Infants' perception of the intonation of broad and narrow focus. *Language Learning and Development, 12*, 1–13.

Chen, A., & Kager, R. (2015). Discrimination of lexical tones in the first year of life. *Infant and Child Development, 25*, 426–439.

Chen, A., Liu, L., & Kager, R. W. J. (2015). Cross-linguistic perception of Mandarin tone sandhi. *Language Sciences, 48*, 62–69.

Chen, A., Liu, L., & Kager, R. W. J. (2016). Cross-domain correlation in pitch perception: The influence of native language. *Language, Cognition, and Neuroscience, 31*, 751–760.

Conboy, B., Brooks, R., Taylor, R., Meltzoff, A., & Kuhl, P. (2008, March). *Joint engagement with language tutors predicts brain and behavioral responses to second-language phonetic stimuli*. Poster presented at the 16th International Conference on Infant Studies, Vancouver, British Columbia, Canada.

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 24–39.

Dietrich, C., Swingley, D., & Werker, J. F. (2007). Native language governs interpretation of salient speech sound differences at 18 months. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 16027–16031.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition, 48*, 71–99.

Escudero, P., & Williams, D. (2014). Distributional learning has immediate and long-lasting effects. *Cognition, 133*, 408–413.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.

Flom, R. (2014). Perceptual narrowing: Retrospect and prospect. *Developmental Psychobiology, 56*, 1442–1453.

Flom, R., & Bahrick, L. E. (2007). The development of infant discrimination of affect in multimodal and unimodal stimulation: The role of intersensory redundancy. *Developmental Psychology, 43*, 238–252.

Frota, S., & Vigário, M. (2008, September). *Early intonation in European Portuguese*. Paper presented at Third Conference on Tone and Intonation (TIE3), Lisbon, Portugal.

Frota, S., Butler, J., & Vigário, M. (2014). Infants' perception of intonation: Is it a statement or a question? *Infancy, 19*, 194–213.

Golinkoff, R. M., Can, D. D., Soderstrom, M., & Hirsh-Pasek, K. (2015). (Baby)Talk to me: The social context of infant-directed speech and its effects on early language acquisition. *Current Directions in Psychological Science, 24*, 339–344.

Gómez, R., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy, 7*, 183–206.

Grossmann, T., Striano, T., & Friederici, A. D. (2006). Crossmodal integration of emotional information from face and voice in the infant brain. *Developmental Science, 9*, 309–315.

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

Harrison, P. (2000). Acquiring the phonology of lexical tone in infancy. *Lingua, 110*, 581–616.

Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development, 65*(3, Serial No. 262).

Huang, T., & Johnson, K. (2010). Language specificity in speech perception: Perception of Mandarin tones by native and nonnative listeners. *Phonetica, 67*, 243–267.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development, 64*, 675–687.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83*, B35–B42.

Krogh, L., Vlach, H. A., & Johnson, S. P. (2013). Statistical learning across development: Flexible yet constrained. *Frontiers in Psychology, 3*. http://dx.doi.org/10.3389/fpsyg.2012.00598.

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

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science, 9*, F13–F21.

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

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

Lany, J., & Saffran, J. R. (2011). Interactions between statistical and semantic information in infant language development. *Developmental Science, 14*, 1207–1219.

Liu, L., & Kager, R. (2011). Is perceptual reorganization affected by statistical learning? Dutch infants' sensitivity to lexical tones. In *Proceedings of the 35th annual Boston University conference on language development* (pp. 404–413). Somerville, MA: Cascadilla Press.

Liu, L. & Kager, R. W. J. (2011). How do statistical learning and perceptual reorganization alter Dutch infants' perception to lexical tones? In W-S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 1270–1273). Hong Kong.

Liu, L., & Kager, R. W. J. (2017). Perception of tones by bilingual infants learning non-tone languages. *Bilingualism: Language and Cognition, 20*, 561–575.

Liu, L., Chen, A., & Kager, R. W. J. (in press). Tone perception in Mandarin and Dutch adult listeners. *Language and Linguistics, 18* (4).

Liu, L., & Kager, R. W. J. (2014). Perception of tones by infants learning a non-tone language. *Cognition, 133*, 385–394.

Liu, L., & Kager, R. W. J. (2013). How bilingualism alters infants' tone perception during perceptual reorganization. In S. Baiz, N. Goldman, & R. Hawkes (Eds.), *Proceedings of the 37th Annual Boston University Conference on Language Development* (pp. 231–240). Somerville, MA: Cascadilla Press.

Männel, C., & Friederici, A. D. (2009). Pauses and intonational phrasing: ERP studies in 5-month-old German infants and adults. *Journal of Cognitive Neuroscience, 21*, 1988–2006.

Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy, 10*, 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*, 1367–1381.

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

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

McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science, 317*, 631.

McMurray, B., Aslin, R. N., & Toscano, J. C. (2009). Statistical learning of phonetic categories: Insights from a computational approach. *Developmental Science, 12*, 369–378.

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

Nazzi, T., Nelson, D. G., Jusczyk, P. W., & Jusczyk, A. M. (2000). Six-month-olds' detection of clauses embedded in continuous speech: Effects of prosodic well-formedness. *Infancy, 1*, 123–147.

Nelson, N. L., & Russell, J. A. (2011). Preschoolers' use of dynamic facial, bodily, and vocal cues to emotion. *Journal of Experimental Child Psychology, 110*, 52–61.

Newport, E. L. (1988). Constraints on learning and their role in language acquisition: Studies of the acquisition of American Sign Language. *Language Sciences, 10*, 147–172.

Ong, J. H., Burnham, D., & Escudero, P. (2015). Distributional learning of lexical tones: A comparison of attended vs. unattended listening. *PLoS ONE, 10*(7), e0133446.

Pannekamp, A., Weber, C., & Friederici, A. D. (2006). Prosodic processing at the sentence level in infants. *NeuroReport, 17*, 675–678.

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

Potter, C. E., Wang, T., & Saffran, J. R. (2017). Second language experience facilitates statistical learning of novel linguistic materials. *Cognitive science, 41*(S4), 913–927.

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

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

Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition, 69*, 105–134.

Seidl, A., & Cristià, A. (2008). Developmental changes in the weighting of prosodic cues. *Developmental Science, 11*, 596–606.

Shukla, M., Nespor, M., & Mehler, J. (2007). An interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology, 54*, 1–32.

Shukla, M., White, K. S., & Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 6038–6043.

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., Nestor, S., Parikh, C., & Yull, A. (2009). Influences of infant-directed speech on early word recognition. *Infancy, 14*(6), 654–666.

Sundara, M., Molnar, M., & Frota, S. (2015). The perception of boundary tones in infancy. In Scottish Consortium for ICPhS 2015 (Eds.), *Proceedings of the 18th International Congress of Phonetic Sciences.* Glasgow, Scotland, UK.

Teinonen, T., Aslin, R. N., Alku, P., & Csibra, G. (2008). Visual speech contributes to phonetic learning in 6-month-old infants. *Cognition, 108*, 850–855.

Thiessen, E. D., & Erickson, L. C. (2012). Discovering words in fluent speech: The contribution of two kinds of statistical information. *Frontiers in Psychology, 3*. http://dx.doi.org/10.3389/fpsyg.2012.00590.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology, 39*, 706–716.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General, 134*, 552–564.

Vanrell, M. M., Prieto, P., Astruc, L., Payne, E., & Post, B. (2010). Early acquisition of F0 alignment and scaling patterns in Catalan and Spanish. In M. Hasegawa-Johnson (Ed.), *Proceedings of Speech Prosody 2010.* Chicago. Retrieved from <http://www.isle.illinois.edu/speechprosody2010/papers/100839.pdf>.

Veenker, T. J. G. (2007). ZEP [computer program]. <https://www.beexy.nl/zep/wiki/doku.php>.

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

Wanrooij, K., Boersma, P., & Van Zuijen, T. (2014). Fast phonetic learning occurs already in 2- to 3-month-old infants: An ERP study. *Frontiers in Psychology, 5*. http://dx.doi.org/10.3389/fpsyg.2014.00077.

Werker, J. F., & Hensch, T. K. (2015). Critical periods in speech perception: New directions. *Annual Review of Psychology, 66*, 173–196.

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

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

Yeung, H. H., Chen, K. H., & Werker, J. F. (2013). When does native language input reorganize phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language, 68*, 123–139.

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

Yeung, H. H., & Werker, J. F. (2009). Learning words' sounds before learning how words sound: 9-month-old infants use distinct objects as cues to categorize speech information. *Cognition, 113*, 234–243.

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