

The more, the better? Behavioral and neural correlates of frequent and infrequent vowel exposure

Sho Tsuji^{1,2,3}  | Paula Fikkert² | Yasuyo Minagawa⁴ | Emmanuel Dupoux⁵ | Luca Filippin⁵ | Maarten Versteegh⁵ | Peter Hagoort⁶ | Alejandrina Cristia^{5,6}

¹ Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania

² Radboud University Nijmegen, Nijmegen, The Netherlands

³ International Max Planck Research School for Language Sciences, Nijmegen, The Netherlands

⁴ Keio University, Tokyo, Japan

⁵ LSCP, Département d'études cognitives, ENS, EHESS, CNRS, PSL Research University, Paris, France

⁶ Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Correspondence

Sho Tsuji, Department of Psychology, University of Pennsylvania, Philadelphia, PA. Email: tsujish@gmail.com

Funding information

Max-Planck-Gesellschaft, International Max Planck Research School for Language Sciences; Region Ile de France; Agence Nationale de la Recherche, Grant numbers: ANR-09-BLAN-0327 SOCODEV, ANR-10-IDEX-0001-02 PSL*, ANR-10-LABX-0087 IEC; Fondation de France

Abstract

A central assumption in the perceptual attunement literature holds that exposure to a speech sound contrast leads to improvement in native speech sound processing. However, whether the *amount* of exposure matters for this process has not been put to a direct test. We elucidated indicators of frequency-dependent perceptual attunement by comparing 5–8-month-old Dutch infants' discrimination of tokens containing a highly frequent [hit-he:t] and a highly infrequent [hyt-hø:t] native vowel contrast as well as a non-native [hɛt-hæt] vowel contrast in a behavioral visual habituation paradigm (Experiment 1). Infants discriminated both native contrasts similarly well, but did not discriminate the non-native contrast. We sought further evidence for subtle differences in the processing of the two native contrasts using near-infrared spectroscopy and a within-participant design (Experiment 2). The neuroimaging data did not provide additional evidence that responses to native contrasts are modulated by frequency of exposure. These results suggest that even large differences in exposure to a native contrast may not directly translate to behavioral and neural indicators of perceptual attunement, raising the possibility that frequency of exposure does not influence improvements in discriminating native contrasts.

KEYWORDS

human infants, language experience, near-infrared spectroscopy, vowel discrimination

1 | INTRODUCTION

During their first year of life, infants' ability to discriminate non-native contrasts declines, while it is maintained or improved for native contrasts (Kuhl, 2004; Tsuji & Cristia, 2013). Frequency of exposure is assumed to play a critical role during this *perceptual attunement*: The more tokens of a given speech sound category infants hear, the more evidence they can accumulate for that particular category in their native language. This assumption is put to a direct test in the present study.

A variety of mechanisms have been proposed for how infants' perception becomes specialized to native categories (e.g., Jusczyk,

1993; Kuhl et al., 2008; Werker & Curtin, 2005). While the field is converging on a scenario where multiple mechanisms interact, such that for instance top-down lexical or referential cues can facilitate speech sound acquisition (e.g., Swingley, 2009; Yeung & Werker, 2009), the assumption that infants profit from the bottom-up accumulation of evidence for a given category nonetheless takes center stage in some theories of speech sound acquisition. Among the proposed bottom-up mechanisms, one possibility is that the perceptual space gets "warped" through the formation of prototypes that act as magnets (Jusczyk, 1993; Kuhl et al., 2008); another is that infants track the frequency of occurrence of acoustic correlates, using

frequency *distributions* rather than prototypes (e.g., Maye, Weiss, & Aslin, 2008; Maye, Werker, & Gerken, 2002). Regardless of the specific mechanism, these proposals assume that frequency of occurrence should play a key role beyond mere presence versus absence of a contrast.

And yet most experimental studies on natural speech have captured developing speech sound perception in a rather categorical way, namely by comparing discrimination of (non-native) contrasts with zero exposure versus (native) contrasts with above-zero exposure. The seminal study by Werker and Tees (1984) was the first to demonstrate how language exposure alters speech sound discrimination during the first year of life, showing that English-learning infants' ability to discriminate two non-native consonant contrasts (a Hindi dental-retroflex contrast [t-t̪], and a Nthlakampx glottalized velar versus uvular contrast [k'-q']) declined between a group aged 6–8 months and a group aged 10–12 months. Hindi- or Nthlakampx-learning 10–12-month-old infants, however, continued to discriminate their respective native contrast. Evidence for perceptual attunement was subsequently also reported with regard to vowel perception. Kuhl, Williams, Lacerda, Stevens, and Lindblom (1992) found that 6-month-old English-learning infants failed to discriminate between prototypical and less prototypical tokens of the native vowel [i], whereas they succeeded in discriminating between non-native tokens of the Swedish vowel [y]. The reverse pattern of discrimination was found in Swedish-learning infants, providing evidence for language-dependent differences in within-category structure. Together with numerous follow-up studies across a large variety of contrasts and languages, these studies show a decline for the discrimination of non-native contrasts, an enhancement for the discrimination of native contrasts, and changes in within-category structure. In addition, researchers have started shedding light on the neural correlates of developing speech sound discrimination. Near-infrared spectroscopy (NIRS) measures changes in blood oxygen level in cortical regions as an index of neural activity. Two measures that have been used in previous work (see, e.g., Minagawa-Kawai, Mori, Naoi, & Kojima, 2007) are the change in blood oxygenation (measured in the superior temporal gyrus, STG, bilaterally), which corresponds to a change detection response, and a laterality index $(L - R)/(L + R)$, which measures the relative left hemisphere advantage and is assumed to reflect increasingly linguistic processing. Broadly speaking, this work shows heightened change detection as well as emergent left dominance for native, but not for non-native contrasts (see Minagawa-Kawai, Cristià, & Dupoux, 2011, for a theoretical review and Tsuji & Cristià, 2013, for an empirical review). For example, Minagawa-Kawai, Naoi, Nishijima, Kojima, and Dupoux (2007) found increased left-hemisphere activation as well as left-dominance at 7–8 months (but not yet at 3–4 months) for the native Japanese contrast between [u] and [o], while no such development was attested for the non-native contrast between [u] and [ʊ].

The above work provides important evidence that the presence versus absence of exposure has an impact on developing speech sound perception. There is only little direct evidence, however, for the key assumption that the *amount* of exposure matters. An

experiment on English infants' discrimination of two non-native contrasts, the coronal Hindi dental-retroflex and the dorsal Nthlakampx glottalized velar-uvular contrast, exploited the fact that coronal [t] is more frequent than dorsal [k] in English (Anderson, Morgan, & White, 2003). While 6.5-month-olds discriminated both non-native contrasts, 8.5-month-olds only discriminated the dorsal one. The authors suggest that frequent native categories become robust earlier, acting as attractors for close non-native contrasts, for which discrimination in turn declines. Pons, Albareda-Castellot, and Sebastián-Gallés (2012) focused on frequency-dependent changes in native discrimination, assessing discrimination of a contrast consisting of one frequent and one infrequent vowel. They showed that both Catalan- and Spanish-learning 12 month olds discriminated [i] and [e] only if the change went from the less frequent to the more frequent speech sound in their respective native language. This perceptual asymmetry indicates that frequent speech sounds can act as attractors to less frequent native speech sounds, actually reducing contrast discrimination in one direction.

These two studies imply that native speech sound frequency influences the decline in non-native discrimination, and that differences in native speech sound frequency lead to asymmetrical discrimination. What remains to be investigated, then, is the impact of input frequency on improvement in native *contrast* discrimination. Specifically, one would expect that the ability to discriminate more frequent native speech sound contrasts (where both speech sounds are frequent) should improve earlier than the ability to discriminate less frequent native speech sound contrasts (where both speech sounds are infrequent). This is the central prediction tested here.

1.1 | The current study

Our central aim was to compare infants' discrimination of a frequent and an infrequent native speech sound contrast to assess the influence of input frequency. We chose two native contrasts that were matched on their respective acoustic distance (as explained in more detail below), and assessed discrimination of the respective contrasts in a between-participants design. In addition, to make sure our experimental design was suited to replicate differences in native and non-native perception, we tested a third group of infants on a non-native contrast that has been attested to show declining discrimination for Dutch infants and adults.

Since individual studies differ in the age at which they report evidence for changes in sensitivity for native vowels (e.g., by 6 months, Kuhl et al., 1992; by 8 months, Minagawa-Kawai, Mori, et al., 2007), we referred to a recent meta-analysis on published studies (Tsuji & Cristià, 2013) to select an appropriate age for observing differences in frequency-related discrimination ability. This analysis showed a significant divergence between native and non-native vowel discrimination after, but not before 6 months of age, thus confirming 6 months as a critical age for perceptual attunement in vowels. We, therefore, decided to test infants in a narrow age range spread around this critical age and to additionally include age as a continuous predictor variable.

The dependent variable in Experiment 1 was looking time differences between non-alternating and alternating trials as an index of discrimination. We predicted a main effect of contrast: Infants should be better at discriminating the native contrasts, than the non-native contrast; moreover, if our predictions are correct, then infants should be better at processing the frequent compared to the infrequent contrast.

2 | EXPERIMENT 1

Data, analysis scripts, and stimuli for this experiment and the next are downloadable from this project's Open Science Framework (OSF) site: <https://osf.io/6vq75/>.

2.1 | Participants

Sixty-two monolingual Dutch full-term infants were included in the final analysis. Twenty-one of these infants were assigned to the frequent condition (8 females, mean age = 6.10 months, range = 4.80–6.80 months), 20 were assigned to the infrequent condition (11 females, mean age = 6.57 months, range = 5.00–7.66 months), and another 21 were assigned to the non-native condition (11 females, mean age: 6.44 months or 196 days; range: 5.29–8.28 months or 161–252 days). Twenty more infants were excluded from data analysis due to fussiness (frequent: 5, infrequent: 2; non-native: 5), failure to habituate (frequent: 2; non-native: 4), experimenter error (infrequent: 1), dialectal language background (infrequent: 1).

2.2 | Stimuli

The Dutch native contrasts were [i – e:] and [ɣ – ø:]. They were selected to differ maximally in their token frequencies, but minimally in their acoustic/perceptual characteristics. As illustrated in Table 1, [i] and [e:] are several times more frequent than [ɣ] and [ø:] (counts based on two corpora of spoken Dutch: CGN, Oostdijk, 2000; IFA corpus, Van Son, Binnenpoorte, van den Heuvel, & Pols, 2001). While the pairs thus differ markedly in frequencies, they have similar acoustic characteristics in that they both consist of a short close vowel and a long diphthongized close-mid vowel. As shown in Figure 1, these tokens are

TABLE 1 Frequency and discriminability of chosen vowels

Contrast	Vowel	Frequency	Discriminability
Frequent	i	10489	0.514
	e:	10087	0.652
Infrequent	ɣ	629	0.413
	ø:	2533	0.661

Frequency counts are token frequencies derived from CGN (Oostdijk, 2000) and IFA corpus (Van Son et al., 2001). Discriminability score are derived from the F1-score (arithmetic mean of precision and recall scores) of the classification of the respective speech sounds in a multi-class classifier model.

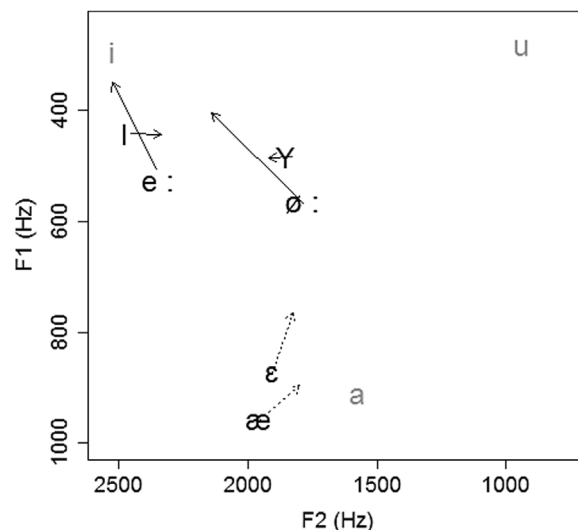


FIGURE 1 Formant values of experimental tokens, and reference point vowels. Endpoints of arrows represent mean formant values in the first and fourth quantiles of the respective vowel (solid lines: native; broken lines: non-native; see also Table 2). Average values for the point vowels [a,i,u] pronounced by 20 female native speakers of Standard Dutch (Adank, van Hout, & Smits, 2004) are added in gray for reference

relatively close in F1/F2 space. In order to measure the similarity of their acoustic/perceptual characteristics, a multi-class classifier model based on mel frequency cepstral coefficients (MFCCs, e.g., Hunt, Lenning, & Mermelstein, 1980) derived from the original tokens in the corpora assessed the discriminability of the four vowels. The discriminability scores of the short vowels are comparable to each other, and the discriminability scores of the diphthongized vowels are comparable to each other (cf. Table 1).

As a non-native contrast, we selected the English vowel pair [ɛ – æ]. This vowel pair has been proven difficult to discriminate by adult native speakers of Dutch (e.g., Broersma, 2002). With regard to Dutch-learning infants, a study testing 6- and 8-month-old Dutch infants with the Hybrid Visual Habituation Procedure found a decline in discrimination ability such that the younger, but not the older age group was able to discriminate this contrast (de Bree, Kerkhoff, de Klerk, Capel, & Wijnen, 2015; see also Wanrooij, Boersma, & Zuijlen, 2014). Note that we did not perform computations on the discriminability of the non-native vowels, since in order to gain a measure that enabled us to compare their discriminability to the native vowels would have required them to be part of the same speech corpus –which is impossible since these vowels do not exist in Dutch. However, the distances as measured in F1/F2 are comparable between native and non-native vowel pairs (see Figure 1 and Table 2). In addition, since the non-native vowel pair served as a control contrast to validate our choice of method and age group, our priority was on choosing a contrast that has been attested to show a pattern of decline for Dutch infants in the literature.

Experimental tokens were recorded in an infant-directed register by a female native speaker of Dutch for the native contrasts, and a

TABLE 2 Acoustic properties of experimental tokens

Contrast	Stimulus	Length (ms)	Mean pitch (Hz)	F1 1st quarter (Hz)	F1 4th quarter (Hz)	F2 1st quarter (Hz)	F2 4th quarter (Hz)
Frequent	hit	411	248	441	444	2456	2335
	he:t	520	255	505	350	2356	2527
Infrequent	hyt	401	262	484	486	1830	1920
	hø:t	521	266	568	404	1786	2142
Non-native	het	420	220	873	763	1904	1826
	hæt	428	220	960	896	1959	1799

Pitch values were averaged over the vowel duration. In order to capture diphthongization, first and second formant values (F1 and F2, respectively) were averaged over the first and last quarter of vowel duration.

female native speaker of American English for the non-native contrasts. The vowels were embedded in a [hVt] context. In preparation for the NIRS study in Experiment 2, for which we closely followed a previous study (Minagawa-Kawai, Mori, et al., 2007) we chose one token of each vowel based on similarity in length and pitch characteristics (cf. Figure 1 and Table 2; waveforms available on OSF project site). For each of the three conditions (frequent, infrequent, non-native), two lists of stimuli, a non-alternating and an alternating list, were created. Non-alternating lists contained 17 repetitions of the stimulus including the short vowel ([hit] in the frequent condition, [hyt] in the infrequent condition, [het] in the non-native condition). Alternating lists contained 16 stimuli in the frequent and infrequent native condition, and 17 stimuli in the non-native condition. Alternating lists repeatedly alternated between the stimulus including the short vowel used in the non-alternating condition and the stimulus including the respective contrasting vowel ([he:t] in the frequent condition, [hø:t] in the infrequent condition, [hæt] in the non-native condition). These lists always started with the stimulus containing the vowel used in the non-alternating condition. The inter-stimulus interval (ISI) was 750 ms for all four lists (Best, McRoberts, & Sithole, 1988). File length for non-alternating lists was 19.79 s for the frequent condition; 19.42 s for the infrequent condition; 19.14 s for the non-native condition. File length for alternating lists was 19.49 s for the frequent condition; 19.61 s for the infrequent condition; and 19.20 s for the non-native condition.

2.3 | Procedure

To assess infants' discrimination abilities, the hybrid visual habituation method (Houston, Horn, Qi, Ting, & Gao, 2007) was implemented with the LOOK software (Meints & Woodford, 2008). Infants were seated in a car seat on their caregiver's lap facing a TV screen. Caregivers were asked not to interact with their infant during the experiment, and both caregiver and experimenter wore headphones with masking music during the course of the experiment. Infants were either assigned to the frequent, infrequent, or non-native condition. The experiment consisted of a habituation and a test phase. Each trial started with a silent attention getter

(a video of a laughing infant). Once the infant looked at the screen, a silent picture of a colorful bull's eye appeared on screen. During habituation, the respective non-alternating list was repeatedly presented to infants until the habituation criterion (50% decrease in looking times compared to the first trials over a sliding window of three trials; see Houston et al., 2007) was reached, or infants had reached a maximum number of 24 trials. During the test phase, infants were presented with 10 non-alternating and 4 alternating trials in pseudo-random order. Three different test orders were created. A trial was terminated when the infant looked away for more than 2 s (Best et al., 1988). In all trials, the visual stimulus was the bull's eye picture. Between trials, the silent movie of a laughing infant appeared to capture infants' attention. The next trial was started once the infant looked at the screen. Looking times were coded online by a trained experimenter.

2.4 | Results and discussion

A linear mixed effect model (lme in nlme; Pinheiro, Bates, DebRoy, Sarkar, & Team, 2012) was fit to the looking time dependent measure, with the between-participant predictor condition (frequent, infrequent, non-native) and age (continuous and centered), the within-participant predictor trial type (non-alternating, alternating), and interactions between these three factors. The intercept and slope were allowed to vary across infants (the precise model was: looking time ~ condition * trial type * age, random = (~trial type|infant), number of observations: 124, number of groups 62). An Analysis of Deviance table was obtained with the Anova function from the car package (Fox & Weisberg, 2011). The model showed a significant main effect of trial type ($\chi^2(1) = 28.85, p < .001$), with higher looking times to alternating ($m = 3.19$ s, $SD = 1.65$) than to non-alternating ($m = 2.41$ s, $SD = 0.79$) trials (Figure 2), and a significant interaction between trial type and condition ($\chi^2(2) = 14.42, p < .001$). No other main effects or interactions reached significance. We followed up on the trial type and condition interaction by constructing separate models with pairwise comparisons between the conditions (p values corrected with false discovery rate). In these models, we found that the interaction between trial type and condition was approaching significance for the

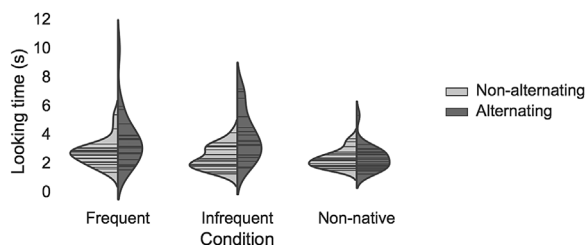


FIGURE 2 Violin plots showing kernel density estimation of the underlying distribution of looking times by Trial Type and Condition in Experiment 1. Horizontal lines represent individual infants' mean looking time

comparison between frequent and non-native ($\chi^2(1) = 4.83, p = .056$), significant between infrequent and non-native ($\chi^2(1) = 17.27, p < .001$) contrasts, and non-significant between the frequent and infrequent contrast ($\chi^2(1) = 2.249, p = 0.134$). Hedges' g effect sizes for the difference in trial type were $g = 0.561$ for the frequent contrast, $g = 1.052$ for the infrequent contrast, and $g = .104$ for the nonnative contrast.

Based on the prediction that the *amount* of exposure matters for perceptual attunement, we hypothesized that the frequent native contrast would be discriminated better compared to the infrequent contrast. However, we found no evidence for this prediction, which is inconsistent with predictions derived from central assumptions in the speech sound acquisition literature. We did, however, find the predicted effect of a divergence in native and non-native perception, suggesting that the experimental design and the chosen age group was, in principle, suited to detect experience-dependent differences in discrimination ability.

3 | EXPERIMENT 2

The behavioral experiment might not have been sensitive enough to detect subtle differences in infant perception elicited by the two native conditions. Infants in the test phase of Experiment 1 looked on average less than 4 s (of a possible 20 s) even to alternating trials, suggesting low attentiveness. Therefore, we turned to NIRS for a more sensitive index of vowel contrast processing. Since auditory change detection in cerebral responses does not depend on the recovery of attention (e.g., Nakano, Watanabe, Homae, & Taga, 2009), NIRS allows for testing both types of contrast in the same session (e.g., Minagawa-Kawai, Mori, et al., 2007). This increases statistical power due to a within-participant design (see Jackson, 2012). In addition, NIRS allows us to observe potential hemispheric differences in processing. A small literature suggests lateralization changes in vowel processing as a function of development, which start even in the first year of infancy (see Minagawa-Kawai, Cristià, & Dupoux, 2011, Figure 1, and Tsuji & Cristià, 2013). As in Experiment 1, we predicted stronger discrimination responses for the frequent than the infrequent contrast.

3.1 | Participants

Thirty-four infants were included in the final analysis (21 females, mean age: 6.71 months, range: 5.45–8.48 months). All infants heard both the frequent and infrequent contrast, with order counterbalanced across infants (18 infants in frequent-first, 16 in infrequent-first). These infants were monolingual Dutch, full-term, in good health and without developmental, language, or hearing problems according to parental report. A further 22 infants were excluded from the analysis for the following reasons: data loss resulting in less than 4 usable trials in each condition: 17; equipment error: 5. Caregivers signed a consent form approved by the local ethical committee (Commissie Mensgebonden Onderzoek Arnhem-Nijmegen, The Netherlands).

3.2 | Stimuli and paradigm

Our experimental design closely followed Minagawa-Kawai, Mori, et al. (2007). The exact same stimuli used in the behavioral study were employed here, with only two changes. First, the interstimulus interval was set to 1.25 s to match previous NIRS studies (Minagawa-Kawai, Mori, et al., 2007). Second, there were two versions of the non-alternating trials, one with 11 repetitions of the non-alternating token (duration 18.79 s), the other with 12 repetitions (duration 20.5 s), this variation in duration serving to jitter the alternating trials. In alternating trials, the two frequent, or the two infrequent tokens were presented in pseudo-random order with equal probabilities every 1.25 s (duration 19.35 or 19.47 s), with a trial always starting with a change token. Unlike Experiment 1, each infant in the present experiment was presented with both conditions, frequent and infrequent (in block design, order counterbalanced across participants). In each condition, infants were presented with non-alternating and alternating trials for a total of eight pairs per condition. The non-alternating trials served as the baseline for the alternating ones.

3.3 | Equipment and data acquisition

Infants were seated on their caregiver's lap in a sound-proof booth and passively listened to the auditory stimuli. An experimenter silently entertained infants with toys during the course of the experiment. Both caregiver and experimenter were wearing headphones with masking music. Stimuli were presented with Psyscope B55 (Bonatti, 2009). The UCL-NTS fNIRS system (Department of Medical Physics and Bioengineering, UCL, London, UK) was used, which continuously emits near-infrared light of two wavelengths, 670 and 850 nm (for further technical details, see Everdell, Coulthard, Crosier, & Keir, 2005; Minagawa-Kawai, Cristià, Vendelin, Cabrol, & Dupoux, 2011). Sources and detectors were positioned on a two by four grid on each of the left and right pads, thus defining a total of 10 channels between optodes separated by 25 mm (see Figure 3), and four more between non-adjacent optodes. In the analyses, we focused on a region of interest (ROI) defined prior to the study in an attempt to match, using anatomical landmarks, the regions tapped by previous work focusing on superior temporal gyrus (Minagawa-Kawai, Mori, et al., 2007):

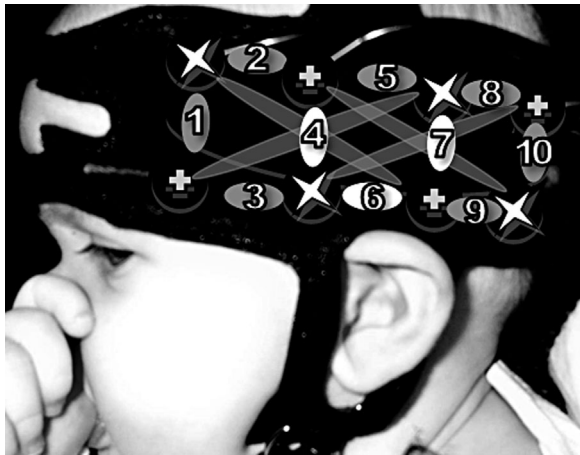


FIGURE 3 Probe array design showing distribution of sources and detectors. Crosses indicate detectors and stars sources. White channels indicate the region of interest

Channels 4, 6, and 7. We used anatomical landmarks to align the bottom of the pad with the T3-T5 line of the 10/20 system, and used the ear as a midpoint reference (see Figure 3). After the data were analyzed, a virtual registration method that does not require having structural neuroimaging data from each participant to estimate the brain regions interrogated by each channel (Tsunami et al., 2007) suggested that the point of maximal sensitivity of channel 6 may be in middle temporal gyrus (rather than the superior one); removing this channel from consideration did not alter the pattern of results described below.

3.4 | Data preprocessing and analysis

Light intensity signals were converted into oxy- and deoxy-Hb concentration with the modified Beer Lambert law. Data were analyzed by applying a general linear model (GLM) to the non-artifacted data of each channel, including nuisance regressors for long-distance, slow trends, and baseline changes as follows (following Cristia et al., 2014; see Minagawa-Kawai, Cristia, et al., 2011 for more extensive descriptions). Slow trends were captured through sine and cosine regressors (for each time-stretch of non-artifacted data longer than 20 s up to the whole duration of the experiment). The data were band-pass filtered between 0.02 and 0.7 Hz only for the following unsupervised artifact detection procedure. Artifacted data were identified as time stretches in which concentration levels changed by more than 0.15 millimolars per millimeter (mM.mm) within 100 ms (time between two successive samples) in the total-Hb averaged over all channels associated with a given probe (Gervain, Macagno, Cogoï, Peña, & Mehler, 2008; Kotilahti et al., 2010). Artifacted stretches were silenced by giving them a weight of zero in the subsequent regression. A boxcar regressor for each new stretch of non-artifacted data were introduced. If there was less than 20 s of unartifacted data between two artifacted regions, this stretch was also silenced, as it is difficult to estimate the hemodynamic response independently from any baseline level changes accompanying an artifact in such short stretches. The data of a channel was altogether

excluded from analysis if there was unartifacted data for fewer than 4 of the total 8 trials in a given condition and infant. Data exclusion did not lead to different number of trials across the two conditions (frequent: $mean = 7.01$, infrequent: $mean = 7.02$), or presentation orders (frequent first: $mean = 7.15$ trials, infrequent first: $mean = 6.85$ trials).

In addition to these nuisance parameters, we declared a regressor based on the standard finite impulse response function (FIR) to estimate specifically concentration changes associated with change detection (again, following Cristia et al., 2014). The first dependent variable, aimed to capture differences in infants' discrimination response, consisted of the beta values obtained from a GLM where the FIR had been convolved with the duration of stimulation for each condition separately. In accordance with previous infant studies, this analysis was based on oxy-Hb (see Lloyd-Fox, Blasi, & Elwell, 2010 for a discussion).

The betas derived from the overall GLM fit were then analyzed with a linear mixed effects model, using the same method as in Experiment 1. In order to assess potential differences in hemispheric lateralization, we included hemisphere as a predictor into the model. Two within-participant predictors, condition (frequent, infrequent) and hemisphere (left, right), and a between-participant predictor, age (continuous, centered), were included as fixed effects together with their interactions, and the intercept and slope were allowed to vary across infants in each condition ($\beta \sim \text{condition} * \text{hemisphere} * \text{age}$, $\text{random} = (\sim \text{condition} | \text{infant})$, number of observations: 349; number of infants 34).

3.5 | Results and discussion

We measured the overall discrimination effect by inspecting the intercept of the full model. The intercept was significant ($t(309) = 3.62$, $p < .001$), indicating that there was an overall difference in processing of non-alternating and alternating trials (cf. Figure 4, left panel). An Analysis of Deviance table revealed no significant effects of any of the predictor variables or their interactions (condition: $\chi^2(1) = 0.683$, $p = 0.409$; hemisphere: $\chi^2(1) = 0.170$, $p = 0.680$; age: $\chi^2(1) = 0.402$, $p = 0.526$; condition \times hemisphere: $\chi^2(1) = 1.252$, $p = 0.263$; condition \times age: $\chi^2(1) = 0.561$, $p = 0.454$; hemisphere \times age: $\chi^2(1) = .004$, $p = 0.948$; condition \times hemisphere \times age: $\chi^2(1) = 1.646$, $p = 0.200$). The average activation change in the left hemisphere was $\beta = .013$ ($SD = .042$, effect size $g = 0.317$) in the frequent condition, and $\beta = .014$ ($SD = .041$, effect size $g = 0.329$) in the infrequent condition. In the right hemisphere, the average activation change was $\beta = .015$ ($SD = .044$, $g = 0.333$) in the frequent condition, and $\beta = .006$ ($SD = .043$, $g = 0.138$) in the infrequent condition.

Since inspection of the curves in Figure 4 suggested to us that there may be slight differences between conditions in infants' brain response in terms of the timing of the hemodynamic response, we followed up on this in additional analyses. These analyses provide some evidence for hemispheric differences in the processing of the contrasts such that the response to frequent stimuli peaked earlier than the response to infrequent stimuli. Since this analysis was, however, a post-hoc analysis potentially leading to false positive

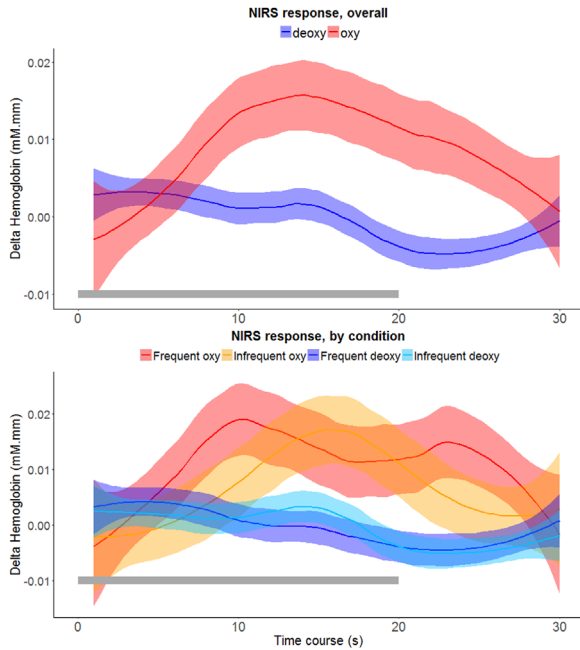


FIGURE 4 Time course of hemodynamic responses, separated by oxy and deoxy Hb (upper panel), and by condition (lower panel). Lines are smoothed with a Gaussian filter only for the purposes of visualization. Shaded areas represent 95% confidence intervals of the average over ROI channels for each infant. The gray bar at the bottom indicates 20 s, the time-window of stimulation

inflation, and since only a little body of studies can be consulted to interpret such a latency difference, these effects will not be further discussed in the present article but are reported in online supporting information in the hope they are replicated in further studies.

Infants' hemodynamic responses (differences in bilateral blood oxygenation) showed the same pattern of results as the behavioral measure in Experiment 1: an overall discrimination effect (suggesting that infants detected the sound change), but no effect of frequency condition, suggesting they were equally successful with frequent and infrequent vowels.

We selected our ROI based on previous research before inspecting our data because this study sought to test a hypothesis and therefore a confirmatory framework was appropriate. To contribute to informing future research endeavors, we addition-

ally illustrate results in a more comprehensive fashion (see Figure 5). For each channel and condition, we conducted a linear mixed effect model [$\beta \sim 1$, $\text{random} = (\sim 1 | \text{infant})$], to report betas reflecting overall response level, and their standard deviation. We refer the interested reader to the dataset on this project's OSF site in case additional calculations are desired. Since we did not find any activations apart from the ROI channels, this descriptive analysis does not conflict in any way with our pre-planned ROI-based one.

4 | GENERAL DISCUSSION

The present study assessed the influence of frequency of exposure on perceptual attunement, and more precisely the prediction that frequently-heard speech sound contrasts should lead to earlier and greater perceptual attunement than infrequently heard contrasts. Given that several models of language acquisition ascribe a central role in perceptual attunement to sounds' frequency of occurrence, documenting such an influence would provide key evidence for the proposed mechanisms.

Overall, however, the data provided no support for these predictions, with both the behavioral and neural measures indicating that infants discriminated both native contrasts equally well, regardless of frequency. How can these findings be interpreted in light of the perceptual attunement literature? One possibility is that there is a difference in perception of frequent versus infrequent contrasts, but at a different point in development, for example, if discrimination ability for both contrasts had already improved prior to testing. This is unlikely, however, since we carefully selected the age range based on a recent meta-analysis of the relevant literature and there is little evidence of narrowing before 5 months of age (see Introduction; Tsuji & Cristia, 2013).

A second possibility is that the chosen contrasts were inherently easy to discriminate for young infants, leading to ceiling discrimination performance. Although the native contrast pairs had a relatively small spectral distance, differences in their duration and dynamics (short vs. long and diphthongized, cf. Figure 1 and Table 2) might have rendered them salient for infants even prior to perceptual attunement. The literature on this question is inconclusive: On the one hand, sensitivity to native vowel length contrasts has been

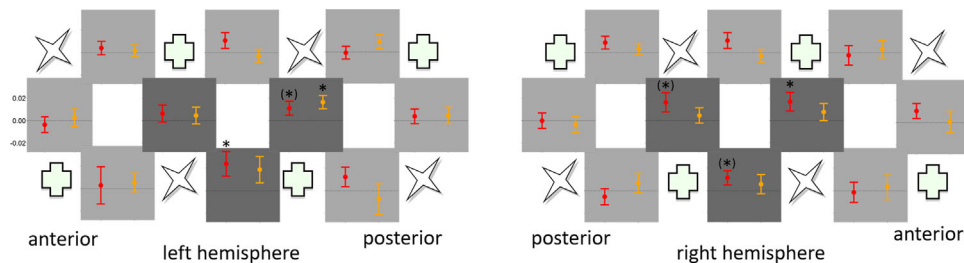


FIGURE 5 Betas from intercepts of overall GLM fits within each channel and condition. Error bars represent $\pm 1SE$, from the model output, and stars denote a significant intercept term (no correction for multiple comparisons applied). Tables with model results can be obtained on the project OSF site

found to develop after the age of 8 months (Sato, Sogabe, & Mazuka, 2010), suggesting the age range we tested should not have been affected by these differences. On the other hand, even 6-month-old infants have been found to discriminate vowel quantity contrasts ([ε-ε:], Pons, Mugitani, Amano, & Werker, 2006) no matter whether they had been exposed to a unimodal or bimodal distribution in a distributional learning setting (Maye et al., 2002, 2008). Thus, it is unclear at this point to what extent the saliency of our contrasts prevented us from seeing a difference between discrimination of frequent and infrequent contrasts, and further studies using different contrasts are desirable to clarify this concern. In this context, it is also noteworthy that the effect size for discrimination of the infrequent contrast was higher than for the frequent contrast in Experiment 1. Since these contrasts were matched as closely as possible in acoustic terms within the available set of native vowels and since no such difference was found in Experiment 2, it is unclear whether this difference indeed reflects a difference in ease of discrimination or would not be replicated in further studies.

Finally, there is a third possibility, namely that frequency of exposure does not matter for improvements in discriminating native contrasts. This option is not unreasonable. For example, learners may require a minimum frequency of occurrence, below which they reduce attention to a salient contrast assuming that it is not present in the native language (as in Anderson et al., 2003), but they would then derive no discrimination benefit from any further exposure above that critical minimum evidence. In other words, one interpretation of our results is that of a true null result, meaning that there is no effect of frequency of occurrence on native contrast discriminability.

Finally, our post-hoc analysis revealed a difference between hemodynamic responses to frequent and infrequent stimuli, albeit not in the strength of discrimination responses as predicted, but in their timing, with faster peak latencies for the frequent condition (see supplementary material). We might be able to observe differences in processing of the present frequent and infrequent contrast at a later stage and on a more linguistic level, for instance in differences in the time-course of left-lateralization (cf. Minagawa-Kawai, Cristià, & Dupoux, 2011). For instance, Japanese infants did not show behavioral differences in their lexical pitch discrimination abilities at 4 and 10 months, but exhibited increased left-lateralization for native pitch contrasts (but not to their pure tone counterparts) by 10 months of age (Sato et al., 2010).

In conclusion, the present study suggests that exposure-dependent differences in infants' vowel processing are not reflected in measures of (behavioral or neural) discrimination or lateralization. The lack of evidence for an influence of frequency of vowel exposure illustrates the need to critically assess the predictions made in the perceptual attunement literature by taking into account central parameters like the nature of the contrast, and the amount of exposure. It has been discussed in several contexts that the saliency of contrasts influences the degree to which perceptual attunement is observable (e.g., Best, 1994; Cristià, McGuire, Seidl, & Francis, 2011; Fava, Hull, & Bortfeld, 2011;

Narayan, Werker, & Beddor, 2010; Werker & Curtin, 2005), but the questions of what determines the saliency of a contrast and how such differences would impact on later linguistic processing still calls for further study. Regarding the amount of exposure, while our study was designed to investigate the effect of relative differences in frequency, an equally or even more critical question might be the absolute amount of input necessary to form a speech sound category. Indeed, our results might reflect that frequency of exposure does not affect the processing of speech sounds once a "critical" number of instances has been encountered, and that the exposure even to very infrequent speech sound contrasts reaches this number relatively early in an infant's life.

ACKNOWLEDGMENTS

We thank Amanda Seidl, Louis ten Bosch, and Christina Bergmann for critical comments during the conceptualization of the study; Christina Bergmann for providing frequency measures for Dutch vowels; Marjolijn van Gelder, Brenda Lelie, Jose Kivits, Imme Lammertink, Lotte Minheere, Charlotte Poulis, and Diana Dimitrova for their assistance in running the study and scheduling infants; Janny Stapel and Sabine Hunnius for their initiative during obtaining ethical approval; Alexander Martin with help in preparing the figures; and the members of the Technical Group of the Donders Centre for Neuroimaging and the Technical Group of the Max Planck Institute for Psycholinguistics for invaluable technical assistance. This work was carried out with a grant of the International Max Planck Research School for Language Sciences; the support of Région Ile-de-France, Agence Nationale pour la Recherche (ANR-09-BLAN-0327 SOCO-DEV; ANR-14-CE30-0003 MechELex) and Fondation de France; and the institutional support from ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL*.

REFERENCES

- Adank, P., Van Hout, R., & Smits, R. (2004). An acoustic description of the vowels of Northern and Southern Standard Dutch. *The Journal of the Acoustical Society of America*, 116, 1729–1738. <https://doi.org/10.1121/1.1779271>
- Anderson, J. L., Morgan, J. L., & White, K. S. (2003). A statistical basis for speech sound discrimination. *Language and Speech*, 46(2–3), 155–182. <https://doi.org/10.1177/00238309030460020601>
- Best, C. T., (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman, & H. C. Nusbaum, (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167–224). Cambridge, MA: The MIT Press, Retrieved from http://www.haskins.yale.edu/SR/SR107/SR107_01.pdf
- Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, 14(3), 345. <https://doi.org/10.1037/0096-1523.14.3.345>
- Bonatti, L. (2009). Psyscope X Build 55 [Computer Program].
- Broersma, M (2002). *Comprehension of non-native speech: inaccurate phoneme processing and activation of lexical competitors*. *Proceedings of the 7th International Conference on Spoken Language*

- Processing. Denver, Colorado: Center for Spoken Language Research, (pp. 261–264).
- Cristià, A., McGuire, G. L., Seidl, A., & Francis, A. L. (2011). Effects of the distribution of acoustic cues on infants' perception of sibilants. *Journal of Phonetics*, 39(3), 388–402. <https://doi.org/10.1016/j.wocn.2011.02.004>
- Cristià, A., Minagawa-Kawai, Y., Egorova, N., Gervain, J., Filippin, L., Cabrol, D., & Dupoux, E. (2014). Neural correlates of infant accent discrimination: An fNIRS study. *Developmental Science*, 17(4), 628–635. <https://doi.org/10.1111/desc.12160>
- de Bree, E., Kerkhoff, A., de Klerk, M., Capel, D., & Wijnen, F. (2015). *Early Speech Perception in Infants with a Familial Risk of Dyslexia*. Philadelphia, USA: Paper presented at 2015 Biennial Meeting of the Society for Research on Child Development.
- Everdell, N. L., Coulthard, M. G., Crosier, J., & Keir, M. J. (2005). A machine for haemodialysing very small infants. *Pediatric Nephrology*, 20, 636–643. <https://doi.org/10.1007/s00467-004-1785-5>
- Fava, E., Hull, R., & Bortfeld, H. (2011). Linking behavioral and neurophysiological indicators of perceptual tuning to language. *Frontiers in Psychology*, 2, -174. <https://doi.org/10.3389/fpsyg.2011.00174>
- Fox, J., & Weisberg, S. An {R} Companion to *Applied Regression*. Second Edition. Thousand Oaks CA: Sage, URL: <http://socserv.socsci.mcmaster.ca/~jfox/Books/Companion>
- Gervain, J., Macagno, F., Cogo, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences of the United States of America*, 105(37), 14222–14227. <https://doi.org/10.1073/pnas.0806530105>
- Houston, D., Horn, D., Qi, R., Ting, J., & Gao, S. (2007). Assessing speech discrimination in individual infants. *Infancy*, 12(2), 119–145. <https://doi.org/10.1111/j.1532-7078.2007.tb00237.x>
- Hunt, M., Lenning, M., & Mermelstein, P. (1980). Experiments in syllable-based recognition of continuous speech. Proceedings of the 1980 International Conference on Audio and Speech Signal Processing, pp. 880–883
- Jackson, S. L. (2012). *Research Methods and Statistics: A Critical Thinking Approach*. 4th edition. Wadsworth: Cengage Learning.
- Jusczyk, P. W. (1993). From general to language-specific capacities: The WRAPSA model of how speech perception develops. *Journal of Phonetics*, 21(1-2), 3–28.
- Kotilahti, K., Nissilä, I., Näsi, T., Lipiäinen, L., Noponen, T., Meriläinen, P., ... Fellman, V. (2010). Hemodynamic responses to speech and music in newborn infants. *Human Brain Mapping*, 31(4), 595–603. <https://doi.org/10.1002/hbm.20890>
- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5(11), 831–843. <https://doi.org/10.1038/nrn1533>
- 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 of London. Series B, Biological Sciences*, 363(1493), 979–1000. <https://doi.org/10.1098/rstb.2007.2154>
- 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(5044), 606–608. <https://doi.org/10.1126/science.1736364>
- Lloyd-Fox, S., Blasi, A., & Elwell, C. E. (2010). Illuminating the developing brain: The past, present and future of functional near infrared spectroscopy. *Neuroscience and Biobehavioral Reviews*, 34(3), 269–284. <https://doi.org/10.1016/j.neubiorev.2009.07.008>
- Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science*, 11(1), 122–134. <https://doi.org/10.1111/j.1467-7687.2007.00653.x>
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82(3), B101–B111. [https://doi.org/10.1016/S0010-0277\(01\)00157-3](https://doi.org/10.1016/S0010-0277(01)00157-3)
- Meints, K., & Woodford, A. (2008). Lincoln Infant Lab Package 1.0: A new programme package for IPL, Preferential Listening, Habituation and Eyetracking. [WWW document: Computer software & manual].
- Minagawa-Kawai, Y., Cristià, A., & Dupoux, E. (2011). Cerebral lateralization and early speech acquisition: A developmental scenario. *Developmental Cognitive Neuroscience*, 1(3), 217–232. <https://doi.org/10.1016/j.dcn.2011.03.005>
- Minagawa-Kawai, Y., Cristià, A., Vendelin, I., Cabrol, D., & Dupoux, E. (2011). Assessing signal-driven mechanisms in neonates: Brain responses to temporally and spectrally different sounds. *Frontiers in Psychology*, 2, 135. <https://doi.org/10.3389/fpsyg.2011.00135>
- Minagawa-Kawai, Y., Mori, K., Naoi, N., & Kojima, S. (2007). Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *The Journal of Neuroscience*, 27(2), 315–321. <https://doi.org/10.1523/JNEUROSCI.1984-06.2007>
- Minagawa-Kawai, Y., Naoi, N., Nishijima, N., Kojima, S., & Dupoux, E. (2007). Developmental changes in cerebral responses to native and non-native vowels: a NIRS study. In Proceedings of the International Conference of Phonetic Sciences XVI (pp. 1877–1880). Saarbrücken. Retrieved from <http://www.icphs2007.de/conference/Papers/1487/1487.pdf>
- Nakano, T., Watanabe, H., Homae, F., & Taga, G. (2009). Prefrontal cortical involvement in young infants' analysis of novelty. *Cerebral Cortex*, 19, 455–463. <https://doi.org/10.1093/cercor/bhn096>
- Narayan, C., 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>
- Oostdijk, N. (2000). Het corpus gesproken nederlands. *Nederlands Taalkunde, Jaargang, 5*, 280–284.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & Team, the R. D. C. (2012). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1–104.
- Pons, F., Albareda-Castellot, B., & Sebastián-Gallés, N. (2012). The interplay between input and initial biases: Asymmetries in vowel perception during the first year of life. *Child Development*, 83(3), 965–976. <https://doi.org/10.1111/j.1467-8624.2012.01740.x>
- Pons, F., Mugitani, R., Amano, S., & Werker, J. F. (2006). Distributional learning in vowel length distinctions by 6-month-old English infants (abstract). Presented at the *International Conference on Infant Studies*, Kyoto, Japan.
- Sato, Y., Sogabe, Y., & Mazuka, R. (2010). Development of hemispheric specialization for lexical pitch-accent in Japanese infants. *Journal of Cognitive Neuroscience*, 22(11), 2503–2513. <https://doi.org/10.1162/jocn.2009.21377>
- Swingle, D. (2009). Contributions of infant word learning to language development. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1536), 3617–3632. <https://doi.org/10.1098/rstb.2009.0107>
- Tsuji, S., & Cristià, A. (2013). Perceptual attunement in vowels: A meta-analysis. *Developmental Psychobiology*, 56(2), 179–191. <https://doi.org/10.1002/dev.21179>
- Tsuzuki, D., Jurcak, V., Singh, A. K., Okamoto, M., Watanabe, E., & Dan, I. (2007). Virtual spatial registration of stand-alone fNIRS data to MNI space. *Neuroimage*, 34(4), 1506–1518.
- Van Son, R. J. J. H., Binnenpoorte, D., van den Heuvel, H., & Pols, L. C. W. (2001). The IFA Corpus: a Phonemically Segmented Dutch "Open Source" Speech Database. Poster presented at Eurospeech (pp. 2051–2054). Retrieved from <http://www.fon.hum.uva.nl/Service/IFACorpus/SLcorpus/AdditionalDocuments/IFACorpusEurospeech2001.html>

- Wanrooij, K., Boersma, P., & Zuijlen, T. L. (2014). Fast phonetic learning occurs already in 2-to-3-month old infants. An ERP study. *Frontiers in Psychology, 5*, 77.
- Werker, J., & Curtin, S. (2005). PRIMIR: a developmental framework of infant speech processing. *Language Learning and Development, 1*(2), 197–234. <https://doi.org/10.1080/15475441.2005.9684216>
- 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*(1), 49–63. [https://doi.org/10.1016/S0163-6383\(84\)80022-3](https://doi.org/10.1016/S0163-6383(84)80022-3)
- Yeung, H. H., & Werker, J. F. (2009). Learning words' sounds before learning how words sound: 9-month-olds use distinct objects as cues to categorize speech information. *Cognition, 113*(2), 234–243. <https://doi.org/10.1016/j.cognition.2009.08.010>

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Tsuji S, Fikkert P, Minagawa Y, et al. The more, the better? Behavioral and neural correlates of frequent and infrequent vowel exposure. *Developmental Psychobiology*. 2017;59:603–612. <https://doi.org/10.1002/dev.21534>