



Lost and Found: Decline and Reemergence of Non-Native Vowel Discrimination in the First Year of Life

Maartje de Klerk, Elise de Bree, Annemarie Kerkhoff & Frank Wijnen

To cite this article: Maartje de Klerk, Elise de Bree, Annemarie Kerkhoff & Frank Wijnen (2019) Lost and Found: Decline and Reemergence of Non-Native Vowel Discrimination in the First Year of Life, *Language Learning and Development*, 15:1, 14-31, DOI: [10.1080/15475441.2018.1497490](https://doi.org/10.1080/15475441.2018.1497490)

To link to this article: <https://doi.org/10.1080/15475441.2018.1497490>



© 2018 The Author(s). Taylor & Francis Group, LLC



Published online: 03 Aug 2018.



Submit your article to this journal [↗](#)



Article views: 513



View related articles [↗](#)



View Crossmark data [↗](#)



Lost and Found: Decline and Reemergence of Non-Native Vowel Discrimination in the First Year of Life

Maartje de Klerk^a, Elise de Bree^b, Annemarie Kerkhoff^a, and Frank Wijnen^a



^aUtrecht institute of Linguistics OTS (Uil OTS), Utrecht University, Utrecht, The Netherlands; ^bResearch Institute of Child Development and Education (RICDE), University of Amsterdam, Amsterdam, The Netherlands

ABSTRACT

Our aim was to investigate perceptual attunement (PA) in vowel perception of Dutch-learning infants (6–8–10-month-olds) using the hybrid visual fixation paradigm (Houston et al., 2007). Infants were habituated to one phoneme and subsequently tested on items in which a token of the habituated phoneme alternated with either another token of the same phoneme, or a token from another phonemic category. Habituation involved tokens of multiple speakers. Infants were tested on a native (/a:-/e:/) and non-native (/ɛ/-/æ/) contrast. The 6-month-olds ($n = 38$), 8-month-olds ($n = 44$) and 10-month-olds ($n = 35$) discriminated the native contrast. The non-native contrast was discriminated by the group of 6-month-olds ($n = 42$) but not the 8-month-olds ($n = 47$), in line with PA. However, the 10-month-olds ($n = 39$) also showed discrimination. We conclude that discrimination of phonetic categories can occur after perceptual attunement; discrimination performance is sensitive to tasks applied.

In acquiring the sound system of their native language, infants learn which acoustic variations indicate phonemic contrasts and which are phonologically irrelevant, such as those resulting from inter- and intra-speaker variation. As a corollary of this learning process, infants' speech perception changes from language-general to language-specific in the first year of life: sensitivity to native speech sound contrasts increases whereas sensitivity to (most) non-native speech sounds decreases (e.g., Cheour et al., 1998; Werker & Tees, 1984). This process is often referred to as perceptual attunement (PA; see Maurer & Werker, 2014 for a recent review). A central prediction of PA is that sensitivity to non-native speech sound contrasts that are assimilated to one native category by adults declines in the first year of life. Although many studies report data that are in agreement with this prediction, not all do (e.g., Best & Faber, 2000; Mazuka, Hasegawa, & Tsuji, 2014; Polka & Bohn, 1996; Tyler, Best, Goldstein, & Antoniou, 2014). Given the lack of uniformity in the literature, further investigation of speech sound discrimination in infancy is warranted. Here, we assess the developmental trajectory of the discrimination of a salient native contrast (serving as a control experiment) and a non-salient non-native contrast in Dutch infants aged six, eight and ten months.

Werker and Tees (1984) were the first to report evidence for PA. They found that English infants discriminated Hindi dental-retroflex plosive (/ʈa/-/ɽa/) and Nthlakampx velar-uvular ejective (/k'i/-/q'i/) contrasts at 6–8 months and 8–10 months, but were not able to do so at 10–12 months of age. In contrast, 11–12-month-old Hindi and Salish learning infants discriminated their native consonant contrasts. Subsequent studies supported PA in consonant perception (e.g., Best, McRoberts, LaFleur, & Silver-Isenstadt, 1995; Sundara, Polka, & Molnar, 2008; Werker & Lalonde, 1988). Although the number of

CONTACT Maartje de Klerk  m.k.a.deklerk@uu.nl  Utrecht institute of Linguistics OTS (Uil OTS), Utrecht University, Trans 10, Utrecht 3512 JK, The Netherlands.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hlld.

© 2018 The Author(s). Taylor & Francis Group, LLC

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

studies on vowel discrimination is limited, findings also show PA, but at an earlier age than for consonants, i.e., around 6–8 months of age (Bosch & Sebastián-Gallés, 2003; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Polka & Werker, 1994; Tsuji & Cristia, 2014).

PA predicts monotonic developmental trajectories. For native speech sounds, early discrimination of highly salient contrasts is anticipated. Early sensitivity is not always found for the less salient native contrasts (Liu & Kager, 2015; Narayan, Werker, & Beddor, 2010; Polka, Colantonio, & Sundara, 2001; Sato, Sogabe, & Mazuka, 2010). When early sensitivity is not attested, gradual acquisition is expected as a result of continued exposure to the native language. This is referred to as enhancement of discrimination or facilitation (Kuhl et al., 2008; Narayan et al., 2010; Tyler et al., 2014). For non-native contrasts, PA predicts a decline in discrimination if the speech sounds of the non-native contrast can be assimilated to one native speech sound. Several studies support this prediction (e.g., Best, 1994; Best & McRoberts, 2003; Cheour et al., 1998; Polka & Werker, 1994; Werker & Lalonde, 1988; Werker & Tees, 1984).

However, a decline in discrimination of non-native speech sounds is not always found (e.g., Best & Faber, 2000; Mazuka et al., 2014; Polka & Bohn, 1996; Tyler et al., 2014). Tyler, and colleagues (2014), for instance, found that English-learning infants at 6 and 11 months discriminated 2 non-native fricative velar-uvular/ χ -/x/ and uvular-pharyngeal/ χ -/ħ/) contrasts from Nuu-Chah-Nulth, a language spoken on Vancouver Island, Canada. PA predicts that the older group would not be able to do so. A similar pattern has been reported for vowel perception. Polka and Bohn (1996) assessed discrimination of a German/u/-/y/ and an English/ ϵ -/æ/ contrast by English- and German-learning infants. They found consistent discrimination by both the German and English-learning infants at 6–8 as well as 10–12 months. Hence, a decline is not always attested.

One factor that might influence discrimination performance is the phoneme used as the standard (habituation) stimulus. Polka and Bohn (1996) found that both English- and German-learning infants showed better discrimination when the habituation stimulus was/y/ than when it was/u/. Similarly, discrimination of the English/ ϵ -/æ/ was attested when the habituation stimulus was/ ϵ /, not when it was/æ/. Polka and Bohn explain these findings with their Natural Referent Vowel framework (NRV; Polka & Bohn, 2011). They propose that vowels in the most peripheral positions of the vowel space (based on their first two formant frequencies, i.e., /i/, a/ and /u/), function as points of reference in the acquisition of the native vowel system. Due to their distinct acoustic and articulatory features, these vowels attract infants' attention (more than non-peripheral vowels do). Consequently, when infants have been habituated to a less peripheral vowel (e.g. German/y/), and are subsequently presented with a peripheral vowel (German/u/), they would show a stronger discrimination response than in the reverse situation. Moreover, the NRV framework proposes that, in general, discrimination is better if the order of presentation is from a less peripheral vowel to a more peripheral vowel, than the reverse. For example, English/æ/ is more peripheral than the English/ ϵ /, and this would explain why discrimination is better when participants are habituated on English/ ϵ / (and subsequently hear English/æ/), than when the habituation stimulus is English/æ/. Although NRV seems to give a plausible explanation for the asymmetries found in speech perception, not all studies assessing discrimination of non-native vowel contrast find discrimination asymmetries that align with the NRV framework (e.g., Best & Faber, 2000; Mazuka et al., 2014; Tyler, Best, Faber, & Levitt, 2014).

Present study

The literature shows that a decline in non-native discrimination over age does not always occur. The present study aims to provide more cross-linguistic data on non-native speech perception. We tracked the development of two types of contrasts: a salient (acoustically and articulatory highly distinctive) native contrast, i.e., the Dutch/a:-/e:/ and a non-salient, non-native contrast, i.e., the English/ ϵ -/æ/ in infants aged six, eight, and ten months old. As the native contrast is salient, we expect that even very young children will be able to discriminate this contrast. Thus, the native contrast serves as a control condition, to assess whether the hybrid visual habituation paradigm

(HVF; Houston, Horn, Qi, Ting, & Gao, 2007) is suitable for assessing speech sound discrimination skills. Selection of a salient native contrast was preferred over a less salient native vowel contrast, such as Dutch/*i/-i:/*: Younger infants might not show evidence of discrimination, as less salient contrasts take longer to acquire (Liu & Kager, 2015). Consequently, using a less salient contrast would not be appropriate for determining the sensitivity of the HVF procedure. To establish whether the sensitivity to a non-native and non-salient (acoustically and articulatory less distinctive) speech sound contrast declines, we chose the English/*ε/-æ/* contrast. We expected that this would be a difficult contrast for the older infants, as (native) Dutch adult listeners assimilate both the English/*ε/* and *æ/* to the Dutch/*ε/* (Broersma & Cutler, 2011; Schouten, 1975).

We used the HVF procedure, which comprises more test trials (fourteen) than traditionally used in speech discrimination research and showed good test-retest reliability (Houston et al., 2007). It is a habituation-dishabituation procedure that combines elements of two other variants of visual fixation procedures. The first is the oddity variant, in which during test the old habituated stimulus is presented less frequently than the new stimulus. The second is the Stimulus Alternation Preference Procedure (SAPP, Best & Jones, 1998), which comprises non-alternating and alternating trials in the test phase. In our study, the procedure starts with habituation to one of the phonemes (e.g., */æ/or/ε/*), and this is followed by a test phase with eight non-alternating (e.g., */æ-æ/or/ε-ε/*) and four alternating pairs (e.g., */ε-æ/or/æ-ε/*)¹.

We used tokens from four different female speakers during habituation. Speaker variability has been argued to enhance generalization of abstract features in the process of developing phonetic categories (Lively, Logan, & Pisoni, 1993; Potter & Saffran, 2015; Rost & McMurray, 2009). In daily speech perception infants need to extract acoustic information that is relevant to phonemic contrasts, while redundant information, not contributing to meaningful differences, needs to be ignored. Hence, the use of multiple speakers makes the task more comparable to the demands of natural speech. Moreover, previous studies assessing discrimination of native and non-native vowel contrasts with multiple speakers have shown that infants are able to extract the relevant acoustic features to distinguish the contrast (Bosch & Sebastián-Gallés, 2003; Sebastián-Gallés & Bosch, 2009).

Two questions were addressed. The first was whether infants discriminate the native contrast at all ages, which we expected to be the case. The second was whether infants show a decline in discrimination performance of the non-native contrast. Here, expectations were less clear-cut. Based on the results of previous studies, both a decline in discrimination and its absence are conceivable (e.g., Polka & Werker, 1994; Polka & Bohn, 1996; Tyler et al., 2014).

Method

Participants

Infants were recruited via the municipality of Utrecht (the Netherlands), and were divided into three age groups: 6– 8- and 10-month-olds. Caregivers were asked to fill out a questionnaire, which asked about birth weight, gestational age, health issues, and family background. Infants were included if: (a) they were raised only in Dutch; (b) their gestational age at birth was considered average, i.e., between 37 and 43 weeks; (c) their birth weight was considered average, i.e., between 2500–5000 grams; (d) there were no complications during the pregnancy or delivery; (e) did not have a history of known hearing loss or reduced vision; and (f) they did not have reported neurological problems.

The aim was to include a minimum of 30 participants who finished both experiments in each age group, divided across habituation stimulus and contrast order. Given the number of anticipated drop-outs this number differs slightly for each age group and contrast (see Table 1, column “Data included”). In total, 366 infants participated; see Table 1 for an overview of the age ranges and drop-out rates per contrast. One hundred and twenty-one infants (33%) were tested but their data was not included in the data analysis. There were different reasons for this: behavior (crying, extreme restlessness, $n = 59$); failure to meet the habituation criterion ($n = 23$; see Procedure); technical errors ($n = 22$); participated in one of the conditions (native or non-native contrast) at an earlier age ($n = 9$, see Table 1); having an

Table 1. Numbers of participants, mean ages and age ranges, and drop-out rates per age group, for each of the two contrasts (native – non-native).

Contrast	Age group	Age range	Age (days)	Data Tested	Drop out 1	Drop out 2	Data Included	Habituation Stimulus 1
		<i>month.days</i>	<i>M (SD)</i>	<i>N</i>	<i>n</i>	<i>n</i>	<i>n (female)</i>	<i>n</i>
Native (/a: /- /e: /)	6	6.1– 6.30	203 (8.4)	59	9	12	38 (18)	17
	8	8.0– 8.30	259 (6.5)	66	8	14	44 (29)	25
	10	10.3– 10.30	320 (12.9)	45	3	7	35 (18)	28
	Subtotal			170	20	33	117 (66)	70
Non-Native (/æ /- /ɛ /)	6	6.1– 6.29	202 (8.2)	65	9	14	42 (12)	25
	8	8.3– 8.29	261 (8.3)	70	9	14	47 (26)	27
	10	10.3– 10.30	325 (6.8)	49	3	7	39 (18)	18
	Subtotal			184	21	35	128 (56)	70
	Total			366	60	61	245 (122)	117

Note. *Drop out 1* refers to infants who did not finish the first experiment or were excluded afterwards. *Drop out 2* refers to infants who did not start or did not finish the second experiment of the session. The column “Habituation Stimulus 1: shows the numbers of participants who were habituated to Stimulus 1, i.e., *faap* in the native condition, and *sæn* in the non-native condition.

ear infection at the time of testing ($n = 5$); parental interference ($n = 2$), or failure to meet the pre- and posttest attention criterion ($n = 1$; see Procedure). Although this drop-out rate is substantial, it can be considered normal for habituation studies (e.g. Narayan et al., 2010; Tyler et al., 2014), especially in a design in which two contrasts were presented subsequently.

Procedure and stimuli

General procedure

The participant was seated on the caregiver’s lap, in a three-walled canvas test booth with a canvas ceiling placed in a sound-attenuated room. The distance between the computer monitor (Philips LCD 150P4) on which the visual stimuli were displayed and the child’s head was approximately 1.35 m. The loudspeaker (Tannoy i8) through which the auditory stimuli were played was hidden behind the canvas of the booth and placed underneath the TV screen that showed the visual stimuli. Caregivers wore headphones (Telex, Echelon 20, over-ear headphones with claimed passive noise attenuation of 20 dB), through which music was played in order to prevent them from hearing the stimuli and (potentially) influencing their child’s behavior. The experiment was monitored and recorded through a video camera that was placed underneath the TV screen.

Caregivers consented to participate during their visit to the lab. In the lab, prior to testing, we explained to the caregiver that two short experiments would be conducted and that the child would hear native and non-native speech sounds, but not in which order this would take place. It was stressed that if a caregiver felt that his/her child was no longer comfortable, they could ask the experimenter to discontinue the experiment at any time. It was also explained that the experimenter could stop the experiment for that same reason. The caregiver was explicitly instructed to (1) not interfere with the experiment, e.g., by pointing to the computer screen, (2) not move their infant during the experimental trials, and (3) soothe his/her child nonverbally when necessary. The aim was to test the infants on both contrasts (native and non-native) within one session. Children with odd numbers were assigned to the native contrast first and the non-native contrast second; children with even numbers were presented with the non-native contrast first and the native contrast second.

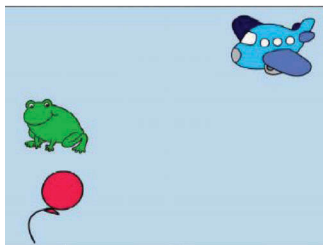
Similar to Houston et al.’s (2007) study, the experiment (both native and non-native conditions) consisted of a habituation phase, in which the infant was habituated on one of the vowels of the pair, a test phase, in which listening times to sequences of trained vowels were compared to those of trained and contrasting vowels, and a pre- and posttest (to measure participants’ attentiveness), in which general listening times were measured. Each of these phases included both auditory as well as visual stimuli.

Stimuli

Auditory and visual stimuli pre-and posttest. During the pre-and posttest infants were presented with both auditory (beep sounds, 330 Hz, duration 250 ms, ISI 1000 ms) and visual stimuli. Auditory stimuli were played at ~ 65 dB(A). The visual stimuli were series of three cartoon pictures (e.g., train, car, book) displayed for two seconds on a light blue screen. These pictures series were drawn randomly from a bank of 25 pictures. They could appear in nine different spots, one per row, within an invisible 3 × 3 grid, see Figure 1. After two seconds new pictures appeared at different locations.

Visual stimuli habituation and test. Visual stimuli were eight still pictures of smiling female faces. Half of these pictures were used during habituation and the other half during test. Pictures were presented in randomized order per block of four trials. Between habituation trials a visual attention getter was displayed: a movie of a cute laughing baby. In between test trials a movie of a toddler going down a slide was used as an attention getter, see also Figure 1.

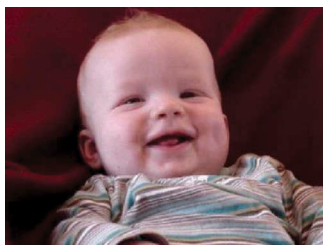
Auditory stimuli habituation and test. Vowel stimuli were presented in CVC syllables (/fa:p/-/fe:p/, /sæn/-/sen/). These targets were pseudowords. Tokens of four different female native speakers were obtained. Auditory stimuli were recorded in a sound-attenuated booth of the phonetics lab of Utrecht University, using a Sennheiser microphone (ME-64) and a digital audio tape recorder (Tascam DA-40). In transferring the recordings onto a computer they were downsampled from 48 kHz to 22.05 kHz. The vowels/a/and/e/were presented in/fVp/-syllables, pseudowords *faap* (/fa:p/) and *feep* (/fe:p/). The four female Dutch speakers were aged between 25 and 35 years of age. They all spoke Standard Dutch and came from the *Randstad* area, a mostly urban area in the central-western Netherlands. Speakers were asked to read out loud a list of 52 words, containing the target pseudowords, as well as other monosyllabic pseudowords and monosyllabic Dutch words with the same vowels (e.g., *gaap* – yawn, *feest* – party). The English [æ] and [ɛ] were presented in/sVn/-syllables, pseudowords/sæn/and/sen/. Tokens were recorded by four female native English speakers, aged between 25 and 35 years. They came from different regions: South-East London, Belfast, Preston (Lancashire) and Manchester. The pseudowords/sæn/and/sen/were read out loud from a list of 52 words containing the target words and real words (e.g., *have* and *pet*) as was done for the native/a:/-e:/contrast.



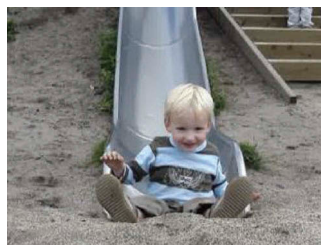
Still of the visual stimuli during pre-and posttest



Picture of smiling female presented during habituation and/or test phase



Still of the attention getter between habituation trials



Still of the attention getter between test trials

Figure 1. Visual stimuli presented during the habituation and test phase.

Each speaker produced four tokens of each target pseudoword (e.g., /fa:p/and/fe:p/). From all four speakers, one token of each target pseudoword per contrast was selected, except from one speaker from whom two tokens per target word were selected. This resulted in five tokens of four different speakers for both contrasts. Four tokens were used during habituation and the fifth token (token 2 from speaker 1, see also [Figure 2](#)) was used in the test phase (see Procedure), hence the fifth token presented during test was from a familiar speaker because participants heard a different token from that speaker during habituation. All auditory stimuli were played at ~ 65 dB(A). Tokens selected were the most child-friendly in prosody and speech affect.

The first and second formant frequencies (Hz) were measured with the software program PRAAT (Boersma & Weenink, 2015, version 5.4.06) and can be found in [Table 2](#). They were measured at the midpoint of the vowel, where the acoustics are minimally influenced by the surrounding consonants. The Dutch tokens are representative for typical/a:/and/e:/vowels spoken by a female as was demonstrated by a study of Adank, van Hout, and Smits (2004). The English recordings of the four speakers had been created to assess categorical perception of these vowels in children and adults (see Heeren, 2006 for a similar approach on/a/-/a:/). The stimuli we used were the end points of these continua. The English tokens were also judged by two native English listeners (from the London area) and rated as good exemplars of the/e/ and/æ/. These tokens are representative for female British-English/e/and/æ/(Deterding, 1997).

Experimental procedure

Pre- and posttest

Pre- and posttest were used to gauge participants' general attentiveness. The pretest started immediately when the participant began to look at the screen and had a fixed duration of approximately 24 seconds. The posttest immediately followed the test phase. Looking times to the screen were measured and were taken to refer to listening times (henceforth listening time). If total listening time to the posttest stimulus was less than 50% of the total listening time to the pretest stimulus, the participant was considered to be showing a general loss of attention. Data of this participant were excluded from further analyses ($n = 1$, see Participants).

Pretest	Habituation Phase	Test Phase	Posttest
Beep sounds 330 Hz 250 ms ISI 1000 ms	Trial 1 /fa:p/ (T1.S1) Trial 2 /fa:p/ (T1.S3) Trial 3 /fa:p/ (T1.S2) Trial 4 /fa:p/ (T1.S4) Trial 5 /fa:p/ (T1.S3) Trial 6 /fa:p/ (T1.S2) Trial 7 /fa:p/ (T1.S4) Trial 8 /fa:p/ (T1.S1) Trial 9 /fa:p/ (T1.S1) Trial 10 /fa:p/ (T1.S2) Trial 11 /fa:p/ (T1.S4) Trial 12 /fa:p/ (T1.S3)	Trial 1 /fa:p/-/fa:p/ (T2.S1 – T1.S1) Trial 2 /fe:p/-/fa:p/ (T1.S1 – T1.S1) Trial 3 /fa:p/-/fa:p/ (T2.S1 – T1.S1) Trial 4 /fa:p/-/fa:p/ (T2.S1 – T1.S1) Trial 5 /fe:p/-/fa:p/ (T1.S1 – T1.S1) Trial 6 /fa:p/-/fa:p/ (T2.S1 – T1.S1) Trial 7 /fa:p/-/fa:p/ (T2.S1 – T1.S1) Trial 8 /fe:p/-/fa:p/ (T1.S1 – T1.S1) Trial 9 /fa:p/-/fa:p/ (T2.S1 – T1.S1) Trial 10 /fa:p/-/fa:p/ (T2.S1 – T1.S1) Trial 11 /fa:p/-/fa:p/ (T2.S1 – T1.S1) Trial 12 /fe:p/-/fa:p/ (T1.S1 – T1.S1)	Beep sounds 330 Hz 250 ms ISI 1000 ms

Figure 2. Schematic of the testing procedure.

Note. In this schematic, the first test trial is non-alternating and the second alternating. The alternative version contains a reversal of these first two trials. In all cases, the remaining three alternating trials have a fixed trial number, namely the 5th, 8th, and 12th trial. Alternating trials are printed in bold. In the habituation phase, speakers are presented in randomized orders per block of four trials. Token is abbreviated and 'T' and speakers as "S".

Table 2. Acoustic characteristics of the stimuli.

Stimulus	Total Duration	Vowel Duration	F1	F2	F3	F0	F0 Range	Intensity
fa:p1	536	210	923	1637	2871	266	201–323	72
fa:p2	587	238	1071	1628	2813	235	185–309	72
fa:p3	559	204	969	1807	2685	291	197–325	72
fa:p4	572	201	941	1615	3033	219	142–331	72
fa:p5	547	205	993	1646	2765	265	198–346	72
fe:p1	570	186	540	2349	2894	267	204–334	72
fe:p2	621	205	618	2168	2851	267	208–308	72
fe:p3	631	191	619	2286	3061	271	196–314	72
fe:p4	622	190	492	2358	3046	210	146–325	75
fe:p5	593	201	513	2383	2873	257	192–334	72
sæn1	556	172	962	1658	3032	221	200–261	74
sæn2	605	228	1036	1863	3337	250	217–309	70
sæn3	484	160	1000	1642	2873	200	186–243	70
sæn4	594	168	1015	1676	2529	212	180–250	71
sæn5	507	181	976	1543	2961	220	200–268	76
sɛn1	459	148	892	2016	3371	245	215–297	75
sɛn2	552	167	760	2061	3340	267	261–293	72
sɛn3	545	163	741	1782	3033	234	225–266	74
sɛn4	665	167	786	2252	3062	222	202–259	72
sɛn5	499	145	889	2005	3315	241	207–285	76

Note. Vowel duration is given in milliseconds. F1–3 refers to the first three formant frequencies, measured at the midpoint of the vowel. F0, the fundamental frequency (pitch), F0 range (minimum – maximum F0) and Intensity are measured over the total duration of the vowel. Stimuli in **bold** are used during test.

Habituation and test

The habituation phase consisted of a maximum of 12 trials, with a maximum of 30 repetitions of a token per trial (ISI of 1 second) resulting in a total duration of approximately 48 seconds. A moving window was used to determine whether the participant had habituated: the mean of trials 1–3 was compared to the mean of trials 4–6. If the mean listening time had decreased with 35%, this was taken as indication that the child had habituated. If listening time had not decreased with 35%, then the mean of the first three trials was compared to the mean listening time of trials 5–7, then 6–8 up to 10–12, as 12 was the maximum number of habituation trials.

The habituation phase started with the attention getter (movie of a cute laughing baby). As soon as the participant looked towards the screen, the experimenter started the first trial. At trial initiation, the visual stimulus changed to one of the smiling female faces, auditory stimuli were played and listening time, was measured. As soon as the participant looked away, the experimenter stopped this measurement and restarted when the infant oriented again to the screen. When the infant looked away for more than two seconds, the trial was terminated and either the next trial started or, if the habituation criterion was reached, the test phase commenced. In the test phase, trials were started and stopped following the same procedure as in the habituation phase. Participants were habituated on either a repetition of /fa:p/ or /fe:p/ tokens. Within one trial, one token of one speaker was used. Participants were presented with all four voices, in randomized order, i.e., in each block of four trials the participant heard all four voices but in randomized order within the blocks. The order of habituation stimuli (*faap* (/fa:p/) or *feep* (/fe:p/)) was counter-balanced between infants.

The test phase had a fixed number of 12 trials, with a maximum number of 30 tokens per trial, resulting in a total duration of approximately 48 sec per trial. Test trials consisted either of alternating pseudoword pairs (i.e., /fa:p/-/fe:p/) or non-alternating pairs (i.e., /fa:p₁/-/fa:p₂/; see *Stimuli*). The alternating and non-alternating trials were presented in a semi-fixed order: the first trial could be either alternating or non-alternating, which was counterbalanced. The second trial was non-alternating if trial 1 was alternating and alternating if trial 1 was non-alternating. Three subsequent alternating trials occurred at positions: 5, 8, and 12. The other trials were non-alternating. During the test phase a new token of a familiar speaker was introduced. This was done to ensure

that the non-alternating trials (e.g./fa:p/-/fa:p/) had both a new token (/fa:p/token-2 from speaker-1) and a familiar token (/fa:p/token-1 from speaker-1), just like the alternating trials had a new token (/fe:p/token-1 from speaker-1) and a familiar token (/fa:p/token-1 from speaker-1) was used. See [Figure 2](#) for a schematic of the procedure.

The test phase started with the attention getter (movie of the toddler on a slide). As soon as the participant looked towards the screen, the experimenter initiated the first trial by pressing a button, which started the trial and listening time measurement. Listening time measurement was the same as during habituation. The changes we made in the design compared to Houston et al.'s study (2007), are summarized in the endnote.¹

Data coding: online and offline

Online coding

The experimenter sat in a room adjacent to the test room and watched the caregiver and infant through a closed-circuit TV. Listening times to trials were captured online by pressing buttons on a button-box connected to a computer (Asus P4PE). An experiment control application (Zep; Veenker, 2008) was used for presentations of the auditory and visual stimuli and for the data registration.

Offline coding

A random subset (approximately 42% of the entire set) of the video recordings was recoded frame-by-frame (one frame had a duration of 30 ms) using Psycode software (<http://psy.ck.sissa.it/PsyCode/PsyCode.html>), by 2 trained coders who were naive regarding to the design and the purpose of the experiment. The results of the raw and recoded data correlated strongly, $r(100) = .99, p < .001$.

Data analysis and screening

Test phase

To answer the questions whether *1) there was an effect of trial type (alternating versus non-alternating), (2) there were differences between the age groups, and (3) the contrasts (native or non-native), the listening times to alternating and non-alternating trials were analyzed using random effect modeling (SPSS, version 23). The raw listening times to alternating and non-alternating trials were not normally distributed; for this reason, a log transformation (Log_{10}) was performed. After this transformation the skewness (.05) and kurtosis (-.37) values were acceptable. Listening times are reported in [Table 3](#) and [Figure 3](#).

Results

The role of contrast

The aim of the study was to investigate the developmental patterns of vowel perception in the first year of life. Our main interest was (1) whether the HVF paradigm could be used to assess the discrimination of speech sound contrasts (rather than word contrasts, as in Houston et al., 2007), and (2) whether non-native discrimination results yielded by the HVF paradigm would agree with PA. Thus, a positive answer to question (1) is a precondition for answering question (2). As infants were tested on both contrasts, we treat *Contrast* (native, non-native) as a within-subject factor. Interactions of *Contrast* with other factors would lead us to analyze the results per contrast separately.

Table 3. Listening times to alternating and non-alternating trials of both contrasts.

Contrast	Age	Alternating Trials		Non-Alternating Trials		Statistics			Participants	Preference for Alternating Trials*	
		<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>	<i>F</i>	<i>p</i>	<i>Cohen's d</i>	<i>N</i>	<i>n</i>	%
Native	6	10.4	(8.6)	7.9	(6.8)	13.55	< .001	.31	38	26	68
	8	9.7	(8.6)	7.1	(6.7)	21.74	< .001	.32	44	30	68
	10	8.1	(5.6)	5.7	(4.5)	29.24	< .001	.45	35	27	77
	All	9.4	(7.9)	7.0	(6.3)	62.70	< .001	.32	117	83	71
Non-Native	6	9.0	(7.7)	7.9	(7.2)	4.59	.032	.15	42	29	69
	8	6.4	(4.7)	6.3	(5.7)	.66	.416		47	25	53
	10	8.6	(8.0)	6.0	(4.3)	21.56	< .001	.37	39	28	72
	All	7.9	(6.8)	6.7	(5.9)	18.16	< .001	.18	128	82	64

Note. Listening times are given in seconds. *Preference for Alternating Trials refers to the number of infants who had on average longer looking times to alternating trials than non-alternating trials.

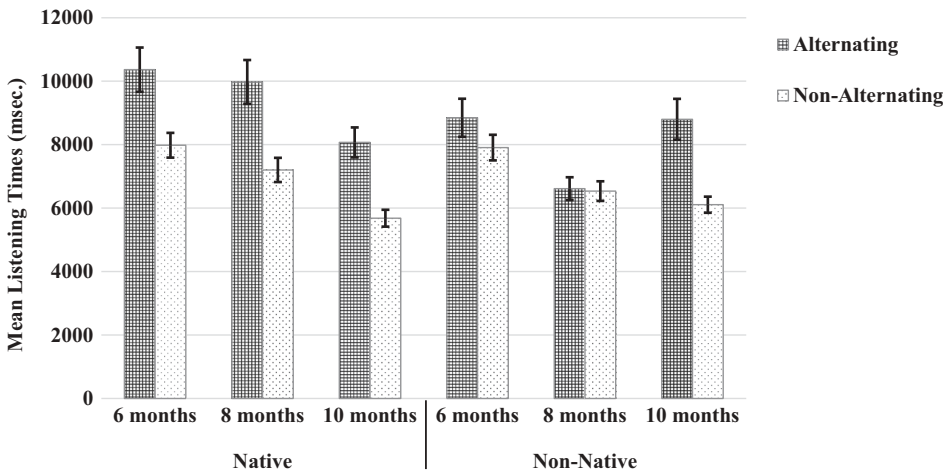


Figure 3. Mean listening times to alternating and non-alternating trials of both contrasts per age group.

Note. Error bars represent standard errors of the mean.

Results of the effect of contrast

Listening times per trial type (alternating vs. non-alternating) are presented in Table 3. A random effect modeling analysis included *Participant* as random factor and *Trial number* as a repeated effect (covariance structure AR1). The fixed factors were *Trial Type* (alternating and non-alternating trials), *Age* (six, eight and 10 months) and *Contrast* (native first vs non-native). The model that best fitted the data included the fixed factors *Trial Type*, *Age*, *Contrast* and *Trial Type*Contrast*Age* $F(8, 650) = 3.05, p = .002$. The 3-way interaction shows that the effect of *Trial type* on listening time differs across contrasts and ages. We will present separate analyses per contrast in the next sections. The main effect of *Trial Type*, $F(1, 1931) = 78.50, p < .001$, indicates that infants listened longer to alternating trials than to non-alternating trials, and the main effect of *Age*, $F(2, 242) = 4.50, p = .012$, means that overall listening time decreased as age increased. No main effect of *Contrast* was found, $F(2, 233) = 1.78, p = .184$, indicating the overall listening times was not significantly different for both contrasts.

The native contrast

Data analysis and screening

Habituation phase

In order to assess whether total listening time and number of trials needed to habituate change as a function of age, univariate ANOVAs and non-parametric tests were conducted. The mean of the total listening times to habituation trials as well as the number of trials required for habituation were assessed across age. The listening times were not normally distributed. Log transformation (Log_{10}) resulted in a distribution that does not differ significantly from a normal distribution (skewness = .09, kurtosis = -.71). The mean number of trials needed to habituate (*NrHab*) was not normally distributed after log transformation. Therefore, non-parametric testing was conducted on this measure.

Test phase

To answer the questions whether (1) infants are able to discriminate the Dutch/a/-/e/contrast, (2) there are differences between the age groups, and (3) the habituation stimulus influences discrimination, the listening times to alternating and non-alternating trials were analyzed using random effect modeling (SPSS, version 23). The raw listening times to alternating and non-alternating trials were not normally distributed; for this reason, a log transformation (Log_{10}) was performed and after this transformation the skewness (.12) and kurtosis (.15) values were acceptable.

Results native contrast

Habituation phase

Although mean listening times showed a tendency to decrease as a function of increasing age (see Table 4), there was no significant main effect of *Age* in a univariate ANOVA with log-transformed mean listening times to habituation trials as dependent variable, $F(2, 109) = 2.49, p = .087$. The number of trials needed to habituate did not differ across age groups (Kruskal-Wallis test, $H(2) = 1.84, p = .912$).

Test phase

Listening times are reported in Table 3 and Figure 3. A random effect modeling analysis included *Participant* as random factor and *Trial number* as a repeated effect (Covariance structure AR1). The fixed factors were *Trial Type* (alternating and non-alternating trials), *Age* (six, eight, and ten months) and *Habituation Stimulus* (/fɑ:p/or/fe:p/). Evidence for continuous discrimination would be visible as a main effect of *Trial Type* and the absence of a significant *Trial Type***Age* interaction. Evidence for directional asymmetry would surface as an interaction between *Habituation Stimulus* and *Trial type*, or in a three-way interaction of *Habituation Stimulus*, *Trial type* and *Age*.

The model that best fitted the data included the fixed factors *Trial Type* (alternating and non-alternating trials) and *Age* (six, eight and 10 months), which comprises a significant effect of 1) *Trial Type* on listening time, $F(1, 916) = 62.59, p < .001$, indicating that infants listened longer to alternating trials than to non-alternating trials, and 2) *Age*, $F(2, 111) = 6.04, p = .003$, meaning that overall listening time decreased as age increased. The *Trial Type* by *Age* interaction was not significant, $F(2, 916) = .90, p = .406$. Nonetheless, planned post hoc comparisons (Bonferroni-adjusted) were conducted to assess whether each age group discriminated the contrast. As can be seen in Table 3 and Figure 3, all age groups discriminated the contrast. Moreover, the effect size of *Trial Type* increases with age, which implies that discrimination becomes more robust as age increases. The models that included the fixed factor *Habituation Stimulus* yielded no effect of *Habituation Stimulus*, $F(1, 111) = .46, p = .500$, no interaction was found between *Trial Type* and *Habituation Stimulus*, $F(1, 926) = .08, p = .783$, and no interaction between *Trial Type*, *Habituation Stimulus* and *Age*, $F(7, 431) = .90, p = .501$.

Table 4. Listening times to habituation stimuli and numbers of habituation trials in the native and non-native contrasts.

Age Group	Contrast	Total Listening Time		Nr. Habituation Trials	
	<i>Native</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
6		108	56	6.6	1.3
8		102	58	6.9	1.6
10		82	42	6.7	1.3
Total		98	54	6.8	1.4
	<i>Non-Native</i>				
6		117	66	6.7	1.6
8		94	47	7.4	1.8
10		83	33	6.8	1.5
Total		101	53	7.0	1.6

Note. Listening times are given in seconds.

Table 5. Numbers of participants, and numbers of participants who received a contrast first and second, per age group and contrast.

Contrast	Age group	Both Contrasts	Contrast Presented 1st	Contrast Presented 2nd
		<i>n</i>	<i>n</i>	<i>n</i>
Native (/a:/-/e:/)	6	31 (0)	15	16
	8	37 (5)	22	15
	10	26 (0)	15	11
	Subtotal	94 (5)	52	42
Non-Native (/æ/-/ɛ/)	6	31 (1)	16	15
	8	37 (3)	15	22
	10	26 (1)	11	15
	Subtotal	94 (5)	42	52
Both Contrasts	Total	188 (10)	94	94

Note. *Both Contrasts* are the number of infants who finished the native and the non-native contrast. The number in parentheses refers to those infants who already heard the contrast at a younger age (at least 2 months earlier) but did not finish the experiment at that earlier session, these infants were not included in the analysis. *Contrast Presented First* are the number of participants who received the contrast first during the session.

Summary native contrast

Our goal was to determine whether the HVF paradigm (Houston et al., 2007) could be used to tap discrimination of speech sounds. For this reason, we chose an acoustically salient (i.e., an acoustically and articulatorily highly distinctive) vowel contrast. The expectation was that infants across the entire age range (six, eight, and ten months) would be able to discriminate the native/a:/-/e:/contrast. This expectation is confirmed by our results. Importantly, this result entails that infants were able to make generalizations over speakers and attend to those acoustic features that differentiate between/a:/and/e:/, regardless of the habituation stimulus. Discrimination performance becomes more robust as age increases, as is indicated by an increasing effect size. These results show that speech sound discrimination in infants can be measured successfully through HVF.

The non-native contrast

Data analysis and screening

Habituation phase

Data analysis was the same as for the native contrast. Listening times to habituation trials were not normally distributed. Log transformation (Log_{10}) rendered a distribution which does not differ significantly from a normal distribution (skewness = .32, kurtosis = -.07).

Test phase

The raw listening times to alternating and non-alternating trials were not normally distributed, but were not significantly different from a normal distribution after log transformation (skewness = .20, kurtosis = -.04).

Results non-native contrast

Habituation phase

Habituation times are reported in Table 4. The numerical decrease of habituation time is not supported by a significant effect of Age on mean (Log_{10}) *Listening Times to Habituation Trials* (Univariate ANOVA), $F(2, 120) = 1.99, p = .141$. The *Number of Trials to Habituate* did also not differ across age groups, $H(2) = 4.39, p = .112$.

Test phase

A change over age in discrimination performance is attested when the interaction *Trial Type*Age* is significant. Whether discrimination is better when trained on one stimulus type would surface as an interaction between *Habituation Stimulus* and *Trial type*, or a three-way interaction between *Habituation Stimulus*, *Trial type* and *Age*. Table 3 and Figure 3 display the results of the test phase.

The model that best fitted the data included the fixed factors *Trial Type* (alternating and non-alternating trials) and *Age* (six, eight, and ten months). The significant *Trial type*Age* interaction $F(2, 1021) = 4.21, p = .015$, was explored by Bonferroni-adjusted pairwise comparisons. Infants aged 8 months did not show a significant difference between alternating and non-alternating trials, whereas the other two age groups did, see Table 3 and Figure 3. As can be seen in Table 3, the effect size of *Trial Type* is larger for the 10-month-olds than for the 6- and 8-month-olds. The effect of *Trial Type* on listening time, $F(1, 1021) = 20.08, p < .001$, indicates that infants listened longer to alternating trials than to non-alternating trials. The effect of *Age*, $F(2, 126) = 2.69, p = .072$, was marginally significant. The pattern points in the direction of a decrease in listening time as age increased, as was found for the native contrast. This finding aligns with results of the total listening time to habituation trials. The models that included the fixed factor *Habituation Stimulus* yielded an effect of *Habituation Stimulus*, $F(1, 121) = 18.15, p < .001$, indicating that infants had overall longer listening times when trained on /sæn/. The interaction between *Trial type*Habituation Stimulus*, $F(1, 1022) = .75, p = .388$, and between *Trial type*Habituation Stimulus*Age*, $F(5, 287) = .56, p = .727$, were not significant; whether infants were trained on either /sæn/ or /sen/ had no influence on discrimination.

The effect of the order in both contrasts

To evaluate whether assessing both native and non-native discrimination within one session impacted infants' performance, we conducted additional analyses in which interactions with *Contrast Order* (first or second) were included (see Figure 4 and Table 5). Again, random effect modeling was used to analyze the data. Fixed factors were *Trial Type* (alternating or non-alternating), *Age* (six, eight, or ten months) *Contrast* (native or non-native) and *Contrast Order* (first or second). The model that best fitted the data included the interaction *Trial Type*Contrast*Age*Contrast Order*, $F(19, 566) = 2.26, p = .002$. The four-way interaction suggests that the effect of *Trial Type* on listening times is not the same for both contrasts at all ages. The interaction between *Trial Type*Contrast Order*Contrast* was marginally significant, $F(5, 727) = 2.09, p = .064$, and suggests that the effect of *Trial Type* on listening times is not the same for both contrast; see Figure 4. It must be noticed, however, that the interactions *Trial Type*Contrast Order* and *Trial Type*Contrast Order*Age* were not significant (all $p < .2$), which means that this four-way interaction should be interpreted with caution.

In order to interpret this four-way interaction we used paired sample t-tests to analyze the data per contrast, contrast order and age group. For the 8-month-olds in the non-native contrast condition we can conclude that no matter whether the contrast was presented first, $t(23) = .18, p = .862$, or second, $t(22) = .25, p = .803$, the contrast was not discriminated. For the 10-month-olds, results are also robust: they

discriminated the contrast when the contrast is presented first, $t(23) = 2.77, p = .011$, and when it is presented second, $t(14) = 2.89, p = .012$. The 6-month-olds, however, seem to discriminate the non-native contrast only when it was presented second, although the effect is only marginally significant, $t(13) = 2.02, p = .064$, and not when it was presented first, $t(27) = 1.12, p = .273$

Order of contrast presentation also affected 6-month olds' performance on the native contrast. They only discriminated the native contrast when it was presented second, $t(16) = 5.04, p < .001$, not first, $t(20) = 1.41, p = .174$. The 8-month-olds discriminate the contrast whether it is presented first (albeit marginally so), $t(28) = 1.99, p = .055$, or second $t(14) = 3.464, p = .004$. For the 10-month-olds the mean difference between alternating and non-alternating trials is only significant if the contrast is presented first, $t(21) = 4.03, p = .001$, not second, $t(11) = 1.76, p = .107$.

From these additional analyses, we conclude that the main findings still hold: 8-month-olds do not discriminate the non-native contrast, whereas the 10-month-olds do.

Summary non-native contrast

The data are suggestive of a decline in non-native/ε/-/æ/discrimination, as predicted by the perceptual attunement hypothesis (e.g., Kuhl et al., 1992; Polka & Werker, 1994; Werker & Tees, 1984): the 6-month-olds discriminated the contrast whereas the 8-month-olds did not. However, the picture is more complex. First, while a significant difference between alternating and non-alternating trials was found for the 6-month-olds, the effect size was small. The claim that 6-month-olds can discriminate the non-native vowel contrast should therefore be made with caution. This is also supported by the additional analyses, which showed that the 6-month-olds did not discriminate the contrast when it was presented first and only marginally so when presented second. Still, this result need not be interpreted as contradictory to PA. Polka & Werker's study (1994) showed that younger infants (4-month-olds) successfully discriminated a non-native vowel contrast, whereas the performance of the 6-month-olds was poorer than predicted. Hence, it is possible that perceptual attunement for vowels starts before or around the age of 6 months.

Secondly, the decline in non-native vowel discrimination was not stable: the 10-month-olds, in contrast to the 8-month-olds, clearly discriminated English/æ/and/ε/. This aligns with other studies which failed to show a decline in discrimination of non-native speech sounds (e.g. Best & Faber, 2000; Polka & Bohn, 1996; Mazuka et al., 2014; Tyler, et al., 2014). In combination with the results of the 6-month-olds, these findings are suggestive of a U-shaped developmental trajectory. A similar pattern is also reported by Best and Faber (2000). They assessed discrimination abilities of English learning infants, aged 3–5, 6–8, and 10–12 months, using a non-native Norwegian (/i/-/y/) contrast with which adult listeners had shown difficulty in an earlier study. The 3–5 and 10–12-month-olds

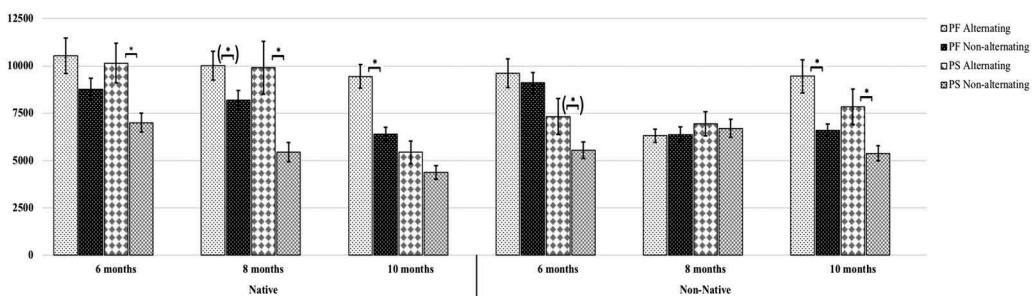


Figure 4. Mean listening times to alternating and non-alternating trials for both contrasts and both orders of presentation.

Note. The abbreviations in the legend, can be read as follows: PF stands for Presented First, which means that the contrast is presented first during the session. PS stands for: Presented Second, which means that the contrast is presented second during the session. Error bars represent standard errors.

did show evidence of discrimination, but the group of 6–8-month-olds did not. The developmental pattern found in the study of Best and Faber thus also shows a “dip” in performance.

General discussion

We aimed to assess whether perceptual attunement occurs in Dutch-learning infants’ vowel perception. Six to ten-month-old infants were tested on a salient native/a:/-e:/contrast and a non-salient, non-native/ε/-æ/contrast. We predicted that the native contrast would be discriminated at all ages, since the contrast we used was a salient (acoustically and articulatorily highly distinctive) contrast. Predictions for the non-native contrast were less straightforward. Based on PA, a decrease in discrimination was to be expected. However, some studies have not found a decline in non-native contrasts (e.g. Best & Faber, 2000; Polka & Bohn, 1996).

The outcome of the first study shows that the HVF paradigm designed by Houston et al. (2007) can be used to assess speech sound discrimination abilities. At all three ages (six, eight, and ten months) infants clearly discriminated the native/a:/-e:/contrast. These results align with earlier findings that salient native contrasts are discriminated by young infants and that this sensitivity is maintained throughout development (e.g., Best et al., 1995; Werker & Tees, 1984).

The findings of the non-native contrast condition are suggestive of a decline in sensitivity between 6 and 8 months of age. This pattern of discrimination performance matches that of PA (e.g., Polka & Werker, 1994). However, in contrast to the PA prediction, our 10-month-old participants showed sensitivity to the non-native vowel contrast. The 10-month-olds discriminated the non-native contrast regardless of whether the contrast was presented first or second. For the 6-month-olds, however, this was not the case. They only discriminated the non-native contrast when it was presented second. The same was found for the 6-month-olds in the native contrast; here too they performed better when it was presented second. These outcomes suggest that the younger infants need some training with the paradigm.

Furthermore, we did not find evidence for discrimination asymmetry (Polka & Bohn, 1996, 2011). Discrimination was not better when children were habituated on/ε/, a less-peripheral vowel, than when they were habituated with the more peripheral/æ/. However, it should be noted that vowel asymmetries are claimed to surface when stimulus presentation changes from the less peripheral vowel to the more peripheral vowel. The HVF procedure might not be suitable to test this, as one vowel type (less or more peripheral) is followed by the other within the same trial. An effect of vowel asymmetry might therefore only be seen in the first non-alternating trial.

It is conceivable that the developmental fluctuations in discrimination attested in this study result from an interaction between the developmental differences between the age groups and the speaker variation used during training. We used multiple exemplars during habituation to facilitate phonetic learning. Variation stimulates phonetic learning as it demands abstraction of invariant features (e.g., Lively et al., 1993). The acoustic variation resulting from speaker variability might have influenced discrimination performance, but in different degrees in each age group. Indeed, there is evidence that the amount of variation needed in order to be helpful during a task differs between age groups (Estes & Lew-Williams, 2015; Singh, 2008; Singh, Morgan, & White, 2004; Vukatana, Graham, Curtin, & Zepeda, 2015). For instance, Singh et al. (2004) showed that 10.5-month-old-infants can recognize a word in a happy affect after having been trained on that same word in a different speech affect (neutral), whereas 7.5-month-olds could not. A follow-up study (Singh, 2008) showed that this latter group did succeed when more variation in speaker affect was offered during training. The amount of variation needed to yield successful (categorical) discrimination seems to vary along age groups. This might explain the U-shaped pattern suggested by our data; the variation may have been enough for the 10-month-olds to support learning, but not for the 8-month-olds.

We argue that the 6-month-olds discriminate the non-native contrast on the basis of their early perceptual abilities, rather than phonetic perception. In this view, the 6-month-olds in our study have not been able to use the speaker variation to discriminate the non-native contrast. The 10-month-olds,

who have acquired native phonetic categories, are able to use their native /e/ category during the experiment: the limited variation offered during training was sufficient for them to make a good estimate of the vowel that was presented during training and maintain a stable representation during test. The 8-month-olds, on the other hand, cannot rely on their early perceptual abilities, nor on their phonetic categories, possibly because they are in the very early stages of PA, i.e., they are in between perceptual strategies. Pursuing this line of reasoning, the amount of variation offered during habituation might not have been sufficient for the 8-month-olds. This leads to the prediction that the 8-month-olds will be able to discriminate the contrast when (much) more speaker variation is introduced. Another prediction is that 10-month-olds will not perform well when there is less variation during training, i.e., when a single speaker is used. We also predict that the amount of variation will not influence the discrimination performance of 6-month-olds. These predictions remain to be tested.

The findings of our study are suggestive of a U-shaped developmental trajectory. Such a pattern has been observed in earlier work investigating the development of native vowel perception of bilingually raised (henceforth bilingual) infants (Bosch & Sebastián-Gallés, 2003; Sebastián-Gallés & Bosch, 2009, but see Burns, Yoshida, Hill, & Werker, 2007; Sebastián-Gallés & Bosch, 2009; Sundara et al., 2008). Bosch and Sebastián-Gallés (2003) tested 4- and 8-month-old Catalan and Spanish monolinguals and Catalan-Spanish bilinguals on a Catalan, but not Spanish, non-salient (acoustically close), /e/-/ε/ contrast (presented in /deði/-/dεði/ pseudowords). They found that Catalan-Spanish bilingual 8-month-olds could not discriminate the contrast, whereas the 4-month-olds could. Their monolingual peers, however, showed the pattern predicted by PA: Monolingual Spanish infants showed a decline in discrimination (as did the bilingual infants), whereas monolingual Catalan infants did not. Subsequently, Bosch and Sebastián-Gallés (2003) tested 12-month-old bilingual Catalan-Spanish infants. This group of infants was able to discriminate the non-salient native /e/-/ε/ contrast. Taken together, the discrimination pattern of the bilinguals over time was U-shaped. In a follow up study with bilingual Catalan-Spanish 6–12-month-old infants (Sebastián-Gallés & Bosch, 2009), however, the U-shaped pattern was only found with another acoustically close (non-salient) contrast /o/-/u/ and not with the salient /e/-/u/. The U-shaped patterns in the studies of Bosch and Sebastián-Gallés (2003, 2009) and our study might both be explained by an interaction between developmental processes such as PA, the salience of the contrast and the experimental design employed.

Indeed, there are indications that the failure of the bilingual Catalan-Spanish 8-month-olds to discriminate the native contrasts is related to the experimental paradigm employed in relation with non-salient stimuli, such as the Catalan /e/-/ε/ contrast. Albareda-Castellot, Pons and Sebastián-Gallés (2011) tested 8-month-old monolingual Catalan and Spanish and bilingual Catalan-Spanish infants on the same vowel contrast /e/-/ε/ as was used in Bosch and Sebastián-Gallés (2003). Instead of a familiarization preference procedure, they used anticipatory eye movement to measure discrimination performance. In their experiment, the performance of the bilingual infants was similar to that of their monolingual peers; both the Catalan monolinguals and the Catalan-Spanish bilinguals discriminated the contrast, while in the study of Bosch and Sebastián-Gallés (2003) the bilinguals failed. The difference between the findings of these two studies might stem from the fact that the familiarization preference paradigm, used in Bosch and Sebastián-Gallés (2003), relies on recovery of attention (increase in listening time) elicited by a vowel change, e.g. a change from /deði/ to /dεði/. However, Albareda-Castellot et al. (2011) indicate that an estimated 66% of all Catalan words have Spanish cognates. Cognates are similar sounding words which often include a vowel difference, e.g., /ʃukulatə/-/tʃokolate/ (*chocolate*). Hence, vowel change does not alter word meaning in many cases and for Catalan-Spanish learning infants, these vowel changes are very common. A paradigm based on the surprise effect of a vowel change might thus not have captured the bilingual 8-month-olds' true sensitivity to this non-salient contrast. So, we argue that the lack of discrimination of the 8-month-old Catalan-Spanish bilinguals in the study of Bosch and Sebastián-Gallés (2003, 2009) is due to the interaction between (1) the contrast being acoustically and articulatory highly similar, (2) PA, and (3) insufficient sensitivity to the paradigm. The lack of discrimination shown in our 8-month-

old monolingual Dutch infants is argued to also be due to type of contrast used (non-salient and non-native), PA and task elements, i.e., insufficient speaker variation during the habituation phase.

In our study, task effects might also explain the large variations in listening times (resulting in small effect sizes) for the 6- and 8-month-olds in native conditions, and the 6-month-olds in the non-native condition. One feature of our procedure that might explain this, is the relatively long ISI (1000 ms). The long ISI might have interfered with the younger groups' discrimination performance, due to their limited short-term memory. Some evidence for this interpretation comes from other studies on vowel perception using a habituation paradigm. Studies that did not find discrimination by very young infants of a non-salient native vowel contrast (Liu & Kager, 2015) or non-native vowel contrasts (Mazuka et al., 2014) had long ISIs (1500 ms). In contrast, a study that did find discrimination by very young infants of a non-salient non-native vowel contrast had a shorter ISI (750 ms, Best & Faber, 2000). Given that working memory capacity increases with age, the effect of shorter ISI duration might be most pronounced at 6 months (Pelphrey et al., 2004). Predictions following from this are that the 6-month-olds will show better performance as a group when ISI is reduced in both native and non-native contrasts.

The results of this study have shown that infants in the process of PA are still able to discriminate a non-native contrast. As Werker (1994, p. 106) states, "developmental changes do not result in a permanent loss" of discrimination abilities. However, during and after the process of PA, discrimination performance might depend to a greater extent on the experimental design.

Notes

- 1 We made six changes to the HVF paradigm as originally described by Houston et al. (2007). First, the test phase was reduced in length, as we know from experience that Dutch children are not always able to sit through experiments that have the same duration as those conducted with children from the U.S. So, instead of 14 test trials, we have 12. The number of alternating trials has remained the same, however.

Second, the target pseudowords were not presented with synchronized audiovisual presentation, as our lab equipment did not allow us to do so. Instead we used still pictures of smiling female faces. Even if the smiles of the smiling female faces interfered with the perception of our CVC pseudowords, they would have affected all vowels equally, since none of the vowels we used are associated with a closed spread position of the mouth; see also Figure 3.

Third, we used multiple speakers in the habituation phase rather than a single speaker. We used multiple speakers to make the task comparable to the demands of natural speech.

Fourth, the habituation criterion was set at 65% instead of 50%. Dijkstra and Fikkert (2010) who used HVF to assess consonant perception, also used the 65% criterion. Other studies assessing speech sound discrimination abilities have also relied on the 65% criterion (e.g., Liu & Kager, 2015; Mazuka et al., 2014; Pater, Stager & Werker, 2004). In our opinion, this criterion allows for tracing a decrease in attention without introducing a risk that infants tune out entirely (which would lead to unwanted data reduction).

Fifth, the pre-test and post-test had a fixed duration. Infants can have very short looking times in the initial phase of an experiment (Colombo & Mitchell, 2009). A fixed duration solves this problem and makes a pre- and posttest with fixed duration a good measure of arousal.

Sixth, we changed the look-away time criterion to 2 sec instead of 1 sec, in light of participants' ages and the stimuli we used. A one second criterion might be too short for the youngest infants to recover from their look away. Many studies assessing speech sound discrimination use the 2-sec criterion (e.g. Best & Faber, 2000; Bosch & Sebastián-Gallés, 2003; Tyler, et al., 2014).

Acknowledgments

We thank the caregivers of our participants for their contributions to our study. We also thank Sule Kurtçebe, Tinka Versteegh, Lorijn Zaadnoordijk and Joleen Zuidema, for assistance in data collection, and Mattis van den Bergh and Kirsten Schutter for their statistical advice.

Funding

This work was supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek [360-70-270].

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. doi:10.1121/1.1779271
- Albareda-Castellot, B., Pons, F., & Sebastián-Gallés, N. (2011). The acquisition of phonetic categories in bilingual infants: New data from an anticipatory eye movement paradigm. *Developmental science*, *14*(2), 395–401. doi:10.1111/j.1467-7687.2010.00989.x
- Best, C., & Faber, A. (2000). *Developmental increase in infants' discrimination of nonnative vowels that adults assimilate to a single native vowel*. Paper presented at the International Conference on Infant Studies (pp. 16, 19), Brighton, UK.
- Best, C., & Jones, C. (1998). Stimulus-alternation preference procedure to test infant speech discrimination. *Infant Behavior and Development*, *21*, 295. doi:10.1016/S0163-6383(98)91508-9
- Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, *167*, 167–224.
- Best, C. T., McRoberts, G. W., LaFleur, R., & Silver-Isenstadt, J. (1995). Divergent developmental patterns for infants' perception of two nonnative consonant contrasts. *Infant Behavior and Development*, *18*(3), 339–350. doi:10.1177/00238309030460020701
- Best, C. C., & McRoberts, G. W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and speech*, *46*(2–3), 183–216. doi:10.1177/00238309030460020701
- Boersma, P., & Weenink, D. (2015) *Praat: Doing phonetics by computer* [computer program]. Version 5. 4.06. Retrieved February, 2015, from <http://www.praat.org>
- Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, *46*(2–3), 217–243. doi:10.1177/00238309030460020801
- Broersma, M., & Cutler, A. (2011). Competition dynamics of second-language listening. *The Quarterly Journal of Experimental Psychology*, *64*(1), 74–95. doi:10.1080/17470218.2010.499174
- 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*(3), 455–474. doi:10.1017/s0142716407070257
- Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature neuroscience*, *1*(5), 351–353. doi:10.1038/1561
- Colombo, J., & Mitchell, D. W. (2009). Infant visual habituation. *Neurobiology of learning and memory*, *92*(2), 225–234. doi:10.1016/j.nlm.2008.06.002
- Deterding, D. (1997). The formants of monophthong vowels in Standard Southern British English pronunciation. *Journal of the International Phonetic Association*, *27*(1–2), 47–55. doi:10.1017/S0025100300005417
- Dijkstra, N., & Fikkert, P. (2010). Do Dutch six-month-olds discriminate *paan* and *taan* and vice versa. Poster session presented at The Boston University Conference on Language Development, Boston, MA.
- Estes, K. G., & Lew-Williams, C. (2015). Listening through voices: Infant statistical word segmentation across multiple speakers. *Developmental Psychology*, *51*(11), 1517. doi:10.1037/a0039725
- Heeren, W. F. L. (2006). *Perceptual development of phoneme contrasts in adults and children* (Doctoral dissertation). Utrecht University. LOT Publications.
- Houston, D. M., Horn, D. L., Qi, R., Ting, J. Y., & Gao, S. (2007). Assessing speech discrimination in individual infants. *Infancy*, *12*(2), 119–145. doi:10.1111/j.1532-7078.2007.tb00237.x
- 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. doi: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 (New York, N.Y.)*, *255*(5044), 606–608. doi:10.1111/j.1467-7687.2006.00468.x
- Liu, L., & Kager, R. (2015). Perception of a native vowel contrast by dutch monolingual and bilingual infants: A bilingual perceptual lead. *International Journal of Bilingualism*, *38*, 27–36. doi:10.1016/j.infbeh.2014.12.004
- Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English/r/and/l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 732–753. doi:10.1121/1.408177
- Maurer, D., & Werker, J. F. (2014). Perceptual narrowing during infancy: A comparison of language and faces. *Developmental Psychobiology*, *56*(2), 154–178. doi:10.1002/dev.21177
- Mazuka, R., Hasegawa, M., & Tsuji, S. (2014). Development of non-native vowel discrimination: Improvement without exposure. *Developmental Psychobiology*, *56*(2), 192–209. doi:10.1002/dev.21193

- Narayan, C. R., Werker, J. F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Developmental Science*, 13(3), 407–420. doi:10.1111/j.1467-7687.2009.00898.x
- Pater, J., Stager, C., & Werker, J. (2004). The perceptual acquisition of phonological contrasts. *Language*, 384–402.
- Pelphrey, K. A., Reznick, J. S., Davis Goldman, B., Sasson, N., Morrow, J., Donahoe, A., & Hodgson, K. (2004). Development of visuospatial short-term memory in the second half of the 1st year. *Developmental psychology*, 40(5), 836. doi:10.1037/0012-1649.40.5.836
- Polka, L., & Bohn, O. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *The Journal of the Acoustical Society of America*, 100, 577. doi:10.1121/1.415884
- Polka, L., & Bohn, O. (2011). Natural referent vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, 39(4), 467–478. doi:10.1016/j.wocn.2010.08.007
- Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d/-/ð/ perception: Evidence for a new developmental pattern. *The Journal of the Acoustical Society of America*, 109(5), 2190–2201. doi:10.1121/1.1362689
- Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 20(2), 421. doi:10.1037/0096-1523.20.2.421
- Potter, C. E., & Saffran, J. R. (2015). The role of experience in children's discrimination of unfamiliar languages. *Frontiers in Psychology*, 6. doi:10.3389/fpsyg.2015.01587
- Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science*, 12(2), 339–349. doi:10.1111/j.1467-7687.2008.00786.x
- Sato, Y., Sogabe, Y., & Mazuka, R. (2010). Discrimination of phonemic vowel length by Japanese infants. *Developmental Psychology*, 46(1), 106. doi:10.1111/j.1467-7687.2008.00786.x
- Schouten, M. E. H. (1975). *Native-language interference in the perception of second-language vowels: An investigation of certain aspects of the acquisition of a second language* (Doctoral dissertation). Utrecht University. LOT Publications.
- Sebastián-Gallés, N., & Bosch, L. (2009). Developmental shift in the discrimination of vowel contrasts in bilingual infants: Is the distributional account all there is to it? *Developmental science*, 12(6), 874–887. doi:10.1111/j.1467-7687.2009.00829.x
- Singh, L. (2008). Influences of high and low variability on infant word recognition. *Cognition*, 106(2), 833–870. doi:10.1016/j.cognition.2007.05.002
- Singh, L., Morgan, J. L., & White, K. S. (2004). Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language*, 51(2), 173–189. doi:10.1016/j.jml.2004.04.004
- Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: Bilingual infants keep pace with their monolingual peers. *Cognition*, 108(1), 232–242. doi:10.1016/j.cognition.2007.12.013
- Tsuji, S., & Cristia, A. (2014). Perceptual attunement in vowels: A meta-analysis. *Developmental Psychobiology*, 56(2), 179–191. doi:10.1002/dev.21179
- Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, 71(1), 4–21. doi:10.1159/000356237
- Tyler, M. D., Best, C. T., Goldstein, L. M., & Antoniou, M. (2014). Investigating the role of articulatory organs and perceptual assimilation of native and non-native fricative place contrasts. *Developmental Psychobiology*, 56(2), 210–227. doi:10.1002/dev.21195
- Veenker, T. J. G. (2008). *The zep experiment control application (version 1.4.1)* [Computer software]. Utrecht Institute of Linguistics OTS, Utrecht University. Retrieved from <http://www.hum.uu.nl/uilots/lab/zep/>
- Vukatana, E., Graham, S. A., Curtin, S., & Zepeda, M. S. (2015). One is not enough: Multiple exemplars facilitate infants' generalizations of novel properties. *Infancy*, 20(5), 548–575. doi:10.1111/infa.12092
- Werker, J. F. (1994). Cross-language speech perception: Development change does not involve loss. In J. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words*. Cambridge, Massachusetts: MIT Press.
- Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology*, 24(5), 672. doi:10.1037/0012-1649.24.5.672
- 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. doi:10.1016/S0163-6383(84)80022-3