Journal of Experimental Child Psychology 206 (2021) 105066

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Speech discrimination in infants at family risk of dyslexia: Group and individual-based analyses

Maartje de Klerk[a,*], Elise de Bree[b], Duco Veen[c], Frank Wijnen[a]

[a] *Utrecht Institute of Linguistics OTS, Utrecht University, 3512 JK Utrecht, the Netherlands*
[b] *Research Institute of Child Development and Education, University of Amsterdam, 1001 NG Amsterdam, the Netherlands*
[c] *Department of Methodology and Statistics, Utrecht University, 3584 CH Utrecht, the Netherlands*

ARTICLE INFO

Article history:
Received 7 June 2020
Revised 10 November 2020
Available online 9 February 2021

Keywords:
Dyslexia
Infants
Speech sound discrimination
Individual analysis
Bayesian hierarchical modeling
Language development

ABSTRACT

Deficiencies in discriminating and identifying speech sounds have been widely attested in individuals with dyslexia as well as in young children at family risk (FR) of dyslexia. A speech perception deficit has been hypothesized to be causally related to reading and spelling difficulties. So far, however, early speech perception of FR infants has not been assessed at different ages within a single experimental design. Furthermore, a combination of group- and individual-based analyses has not been made. In this cross-sectional study, vowel discrimination of 6-, 8-, and 10-month-old Dutch FR infants and their nonrisk (no-FR) peers was assessed. Infants (N = 196) were tested on a native English /aː/–/eː/ and non-native English /ɛ/–/æ/ contrast using a hybrid visual habituation paradigm. Frequentist analyses were used to interpret group differences. Bayesian hierarchical modeling was used to classify individuals as speech sound discriminators. FR and no-FR infants discriminated the *native* contrast at all ages. However, individual classification of the no-FR infants suggests improved discrimination with age, but not for the FR infants. No-FR infants discriminated the *non-native* contrast at 6 and 10 months, but not at 8 months. FR infants did not show evidence of discriminating the contrast at any of the ages, with 0% being classified as discriminators. The group- and individual-based data are complementary and together point toward speech perception differences between the groups. The findings also indicate that conducting individual analyses on hybrid visual habituation outcomes is possible. These outcomes form a fruitful avenue for gaining more understanding of development, group differences, and prospective relationships.

---

[*] Corresponding author.
E-mail address: m.k.a.deklerk@uu.nl (M. de Klerk).

https://doi.org/10.1016/j.jecp.2020.105066

Introduction

Developmental dyslexia is a language-based learning disability characterized by severe word reading and/or spelling problems (Lyon, Shaywitz, & Shaywitz, 2003; Peterson & Pennington, 2015). These literacy difficulties can have a profound impact on educational/academic achievement, self-esteem, and social development (Livingston, Siegel, & Ribary, 2018). Therefore, it is of great interest to understand the precursors or risk factors that lead to the subsequent deficit. Dyslexia is considered to be a multifactorial disorder, which implies that multiple risk and protective factors are involved. The disorder is considered to be highly heritable; children with a dyslexic parent have a 29% to 66% risk of developing dyslexia (Snowling & Melby-Lervåg, 2016). Assessing the abilities of children with a family risk (FR) of dyslexia, therefore, is a valuable approach for finding early markers of dyslexia (e.g., Caglar-Ryeng, Eklund, & Nergård-Nilssen, 2019; Snowling & Melby-Lervåg, 2016; van Viersen et al., 2018).

A phonological deficit has been proposed to be one of the main contributing risk factors in developing dyslexia (Ramus et al., 2003; Vellutino, Fletcher, Snowling, & Scanlon, 2004). It is hypothesized that people with dyslexia have poorly specified phonological representations, which in turn have a disruptive effect on the construction of phoneme–grapheme connections (e.g., Blomert, 2011; Mittag, Thesleff, Laasonen, & Kujala, 2013). Although the phonological deficit cannot account for all literacy problems in people with dyslexia (Pennington et al., 2012), there is extensive evidence for phonological problems in children and adults diagnosed with dyslexia (Ramus et al., 2003; van Bergen, de Jong, Plakas, Maassen, & van der Leij, 2012). It is also seen in FR children prior to the acquisition of literacy skills (see Snowling & Melby-Lervåg, 2016, for a meta-analysis and review), which is suggestive of a causal relation.

One potential cause of the phonological deficit is poor speech perception; if speech sounds cannot be perceived and categorized adequately, this will hamper the formation of phonological representations as well as grapheme–phoneme associations (e.g., Goswami, 2000). A large number of studies found that adults and children with dyslexia perform more poorly on tasks measuring speech perception skills than their peers (e.g., Schulte-Körne, Deimel, Bartling, & Remschmidt, 2001; Werker & Tees, 1987; but see Nittrouer, Shune, & Lowenstein, 2011; Ramus et al., 2003; Rosen & Manganari, 2001). This is also found for children and infants with an FR (e.g., Boets, Ghesquière, Van Wieringen, & Wouters, 2007; Guttorm et al., 2005; Richardson, Leppänen, Leiwo, & Lyytinen, 2003; van Alphen et al., 2004). Moreover, some studies have found that children with lower (pre)reading skills showed poorer speech perception performance as infants (e.g., Guttorm, Leppänen, Hämäläinen, Eklund, & Lyytinen, 2010; Molfese, 2000; van Zuijen, Plakas, Maassen, Maurits, & van der Leij, 2013). Hence, speech perception skills seem to be related to learning to read and spell effectively and efficiently.

Adults and children with dyslexia have been found to perform more poorly than their peers on speech sound categorization tasks (Hakvoort et al., 2016; Maassen, Groenen, Crul, Assman-Hulsmans, & Gabreëls, 2001; Mody, Student-Kennedy, & Brady, 1997), which has led to the proposal of a categorical speech perception deficit (e.g., Serniclaes, van Heghe, Mousty, Carré, & Sprenger-Charolles, 2004). One explanation for the categorization deficit could be that "children with dyslexia maintain the sensitivity to phonemic distinctions which all newborns have irrelevant of their native language" (Noordenbos, Segers, Serniclaes, Mitterer, & Verhoeven, 2012, p. 1470). Although the reported results are not fully consistent (Blomert & Mitterer, 2004; Brandt & Rosen, 1980; Hazan, Messaoud-Galusi, Rosen, Nouwens, & Shakespeare, 2009; Messaoud-Galusi, Hazan, & Rosen, 2011), a recent meta-analysis did show support for a categorical perception deficit in dyslexia (Noordenbos & Serniclaes, 2015). Poor categorical perception has also been found in kindergartners with an FR of dyslexia (Boets et al., 2007; Gerrits & de Bree, 2009; Noordenbos et al., 2012).

Phonological categorization builds on a robust speech sound discrimination ability. The available evidence suggests that speech sound discrimination in FR infants is weaker in comparison with low-risk (no-FR) peers (van Leeuwen et al., 2006; see Richardson et al., 2003, and Volkmer & Schulte-Körne, 2018, for recent reviews on electroencephalography [EEG] studies). van Leeuwen et al. (2006) conducted an EEG study with 2-month-old Dutch infants using an oddball paradigm in which /b/ and /d/ were presented in Dutch /bak/ (*box*)–/dak/ (*roof*) words. The tokens used were taken from a /b/–/d/ continuum. The FR infants showed a significantly less pronounced mismatch negativity response to the deviant stimulus, indicative of a delay in categorization. Such poorer phoneme discrimination in FR infants has been found for consonants as well as vowels (e.g., Guttorm, Leppänen, Richardson, & Lyytinen, 2001; Leppänen, Pihko, Eklund, & Lyytinen, 1999; Molfese, 2000; Pihko et al., 1999; Thiede et al., 2019; van Leeuwen et al., 2006). In sum, the literature shows that FR infants have more difficulty with discrimination between phonemes. This finding can be related to subsequent poor categorization and aligns with the notion of a speech perception deficit in dyslexia.

The studies that report discrimination difficulties in FR infants have so far been limited in the sense that all the speech sound contrasts under investigation were native contrasts and were mostly assessed at one age. Because it is well established that speech perception changes during the first year of life due to language exposure, it is warranted to investigate how (native) speech perception develops in FR infants. In typically developing infants, speech perception changes from universal to language specific (e.g., Werker & Tees, 1984). This means that the ability to discriminate native speech sound categories remains good or improves (for sounds that are initially difficult to discriminate), whereas the ability to detect speech sound distinctions that are not phonemic in the native language decreases (e.g., Tsuji & Cristia, 2014; Werker & Tees, 1984). This developmental transition is generally referred to as *perceptual attunement* (Maurer & Werker, 2014) and emerges around 10 to 12 months of age for consonantal contrasts and at 6 to 8 months for vowel contrasts (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Polka & Werker, 1994). Perceptual attunement is the first step in the formation of (native) phoneme categories.

Although investigation of the developmental trajectory of native and non-native speech perception of FR infants is important for evaluating the process of perceptual attunement in FR, we know of no studies that have looked into this. There are, in contrast, some studies with no-FR infants. In one such study (de Klerk, de Bree, Kerkhoff & Wijnen, 2019), it was found that no-FR infants were able to discriminate between the salient native vowel contrast /aː/ and /eː/ at 6, 8, and 10 months of age and that this discrimination improved with age. In contrast, only the 6-and 10-month-olds were able to discriminate between non-native English /ɛ/ and /æ/; the 8-month-olds were not. These findings are indicative of perceptual attunement from 6 to 8 months. The finding that the 10-month-olds could discriminate the contrast was explained by an interaction between task demands and maturation (de Klerk et al., 2019).

The current study

The current study compared speech sound discrimination of 6-, 8-, and 10-month-old FR infants with that of their no-FR peers. We used the hybrid visual habituation paradigm (de Klerk et al., 2019; Houston, Horn, Qi, Ting, & Gao, 2007) comprising test trials with similar phonemes (non-alternating, e.g., /aː/–/aː/) and different phonemes (alternating, e.g., /aː/–/eː/). We addressed two questions. The first is whether perceptual attunement occurs in FR infants. In other words, is there a change from a universal listener to a language-specific listener also in FR infants? If perceptual attunement takes place, native contrasts were expected to be discriminated at all ages, whereas there would be a decrease in the ability to discriminate non-native contrasts as infants mature (e.g., de Klerk et al., 2019; Kuhl, et al., 2008; Tsuji & Cristia, 2014). Hence, FR infants should be able to discriminate a salient (acoustically and articulatory highly distinctive) native contrast such as Dutch /aː/–/eː/. Selection of a salient native contrast was preferred over a less salient native vowel contrast, such as Dutch /ɪ/ - /iː/, as such less salient native contrasts take longer to acquire (e.g., Liu & Kager, 2015) and might thus not be discriminated at all by the infants in our sample. Studies on speech sound discrimination skills of FR infants have often used nonsalient native contrasts and found perception difficulties on these subtle contrasts (e.g., van Leeuwen et al., 2006). Therefore, it is possible that initial discrimination is weak(er) and the gradual improvement of discrimination proceeds more slowly than in no-FR infants. In other words, there could be a delay in perceptual attunement for FR infants. Hence, investigation of a native salient contrast is warranted.

With respect to the discrimination of non-native contrasts, the question was whether FR infants' data would provide evidence of discrimination at 6 months of age and a decrease in sensitivity at later ages to the English /ɛ/–/æ/ contrast that we used in this study. This contrast is difficult to distinguish for Dutch adults (Broersma & Cutler, 2011); both vowels are perceived as Dutch vowel /ɛ/. On the basis of previous findings of poor speech sound discrimination of FR infants (e.g., Leppänen et al., 2002; Richardson et al., 2003; van Leeuwen et al., 2006), it is conceivable that the subtle non-native contrast would not be discriminated throughout development (e.g., Richardson, et al., 2003; van Leeuwen et al., 2006). Another possibility was that infants would not lose the sensitivity to the irrelevant non-native contrast (Noordenbos et al., 2012).

Our second main question was whether it is possible to identify individual infants as being able to discriminate the speech sound contrast or not. Put differently, can infants be classified as "discriminators" at the individual level on the basis of outcomes on behavioral speech perception tasks? This is an important question in the field of speech perception (Cristia, Seidl, Junge, Soderstrom, & Hagoort, 2014; Houston et al., 2007): Group-based findings are valuable for understanding a general pattern of discrimination, but identification of individual difficulties and future outcomes requires reliable analyses on individual-based data. Furthermore, studies have investigated the relation between early speech perception and later reading skills retrospectively (Guttorm et al., 2010; Molfese, 2000; van Zuijen et al., 2013). If discriminators at infancy can be identified successfully, this could pave the way for prospective studies into early speech perception and later language and reading skills.

To address this question, we took our previous study (de Klerk, Veen, Wijnen & de Bree, 2020) as a starting point. We evaluated different methods of individual analyses and found that Bayesian hierarchical modeling was the most successful one. This approach takes into account the hierarchical nature of the data; infants within the same age group are assumed to belong to the same population, meaning that infants are exchangeable *within* age groups but not *between* age groups. The advantage of Bayesian hierarchical modeling in comparison with frequentist approaches is that it yields estimates for all the individual and group parameters in one model without needing to correct for multiple testing (Gelman, Hill, & Yajima, 2012). Furthermore, the consequence of hierarchically modeling the individual and group effects in one analysis is that part of the observed variance can be explained at the group level instead of trying to explain all the variance at the individual level (Gelman, 2006). In our previous study (de Klerk, Veen et al., 2019), we showed that by adding the hierarchical structure, we reduced the noise, which led to less variable representations of the measurements. This can be seen as an improvement of the reliability of the measurements (Gelman et al., 2012). Individual outcomes can provide more insight into developmental trajectories and thus are of great value for studies that relate early abilities to later language skills.

Method

Participants

Participants were recruited via a letter sent to all the parents of newborns of Utrecht City, the Netherlands. Addresses were supplied by the municipality of (Utrecht City, the Netherlands). Before coming to the lab, parents were asked to fill out a questionnaire consisting of questions about birth weight, gestational age, health issues, and (medical) family background. Infants were included if (a) they were raised only in Dutch, (b) their gestational age at birth was 37 to 43 weeks, (c) their birth weight was 2500 to 5000 g, (d) there were no complications during the pregnancy or delivery, (e) they did not have a history of known hearing loss or reduced vision, and (f) they did not have reported neurological problems.

To ascertain whether the FR infants could truly be categorized as such, three tests were administered to parents who had indicated a history of reading problems. The first was a timed word reading test, the Een-Minuut-Test (EMT; Brus & Voeten, 1972). In this test, parents needed to read out loud a list of known words as quickly and accurately as possible within 1 min. The second test, de Klepel, was a timed pseudoword reading test. Parents were asked to read out loud a list of pseudowords within 2 min (van den Bos, Lutje Spelberg, Scheepsma, & de Vries, 1994). The third test, a verbal competence test (Analogies), was a subtest of the Dutch version of the Wechsler Adult Intelligence Scale (WAIS; Uterwijk, 2000). Infants were included in the FR group if parents had met one of the following criteria: (a) the percentile scores on one reading test was ≤10, (b) the percentile score was ≤20 on both reading tests, or (c) the discrepancy between one reading test and the verbal competence test was 60 percentile points or more (Kuijpers et al., 2003). If the criteria were not met, the infant was not included in the study (*n* = 7).

A flowchart of the data inclusion process can be found in Fig. 1. In total, 117 FR infants were tested, potentially on both the native and non-native vowel discrimination experiment, rendering a potential of 234 datasets or records (2 * 117 records). However, 71 records (30%) were not included for the following reasons: (a) behavior invalidating the measurements (e.g., crying, extreme restlessness; *n* = 22); (b) the second discrimination experiment was never started (the decision to proceed to the next experiment depended on the behavior and well-being of the infant after the first experiment; *n* = 23); (c) the parent was not classified as dyslexic (*n* = 14 records; see above); (d) failure to meet the habituation criterion (*n* = 11; see "Procedure" section below); or (e) technical error (*n* = 1). Fig. 1 contains a capture of the inclusion criteria in a flowchart. In total, 163 records were included. These 163 records came from 98 infants. Of these, 65 infants finished both the native and non-native conditions. Hence, these 65 infants yielded 130 records (native *n* = 65 and non-native *n* = 65). Some infants (*n* = 33) finished only one contrast (native, *n* = 14 or non-native, *n* = 19). See Table 1 (FR infants) and Appendices A and B.

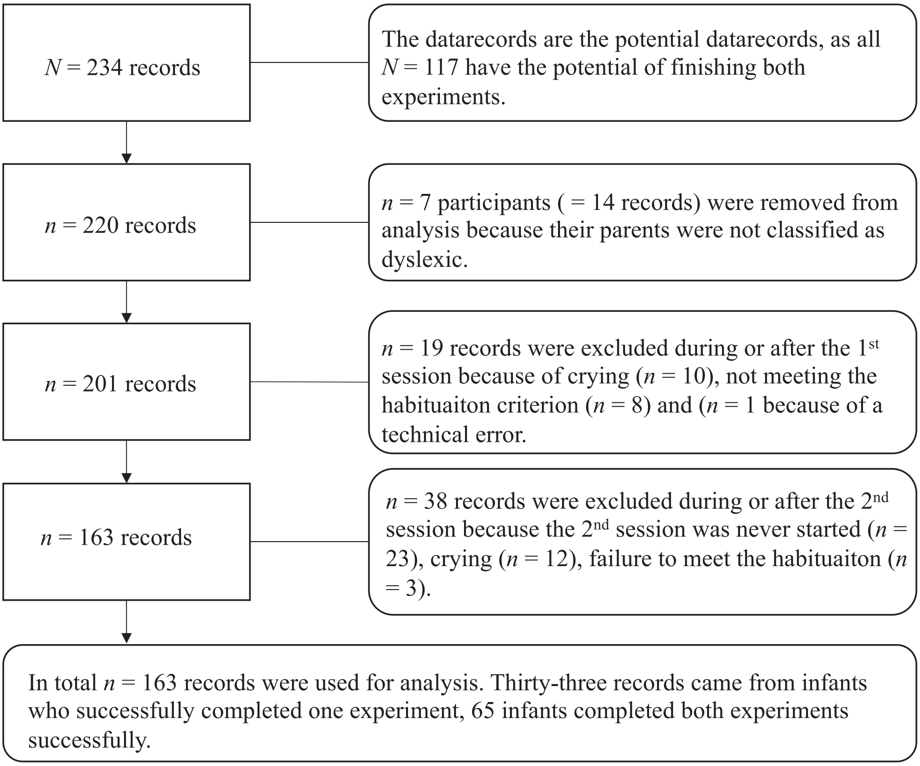


Fig. 1. Flow chart for data inclusion.

The no-FR infants were selected from the data set presented in de Klerk et al., 2019. The no-FR infants (n = 98) were matched to the FR infants on the following characteristics: (a) age, (b) the number of experiments they had finished during the session (1 or 2), (c) the stimulus they were habituated on, and (d) which contrast was presented first (native or non-native). In the no-FR selection as well, 65 infants completed both experiments and 33 infants finished the task in one condition. See Table 2, Appendix A (native contrast), and Appendix B (non-native contrast) for more information regarding the number of infants per age group who finished both contrasts.

All parents were native monolingual speakers of Dutch and lived in Utrecht City, the Netherlands. Data on parental level of education and the family situation (i.e., the number of siblings and birth order [first born, second born, etc.]) are summarized in Table 3. The educational level was coded ranging from 1 (primary school) to 6 (university level). The average educational level of FR infants' fathers was significantly lower than that of no-FR fathers, but for both groups the educational level was high (see Table 3). The majority of FR fathers (88%) and no-FR fathers (95%) had completed a university degree (bachelor's or master's level).

Informed consent was obtained from the parents before testing; consent and participation could be retracted at any time. The research was conducted in accordance with American Psychological Association ethical standards as well as The Netherlands Code of Conduct for Scientific Practice issued in 2004 (revised in 2018 by the Association of Universities in The Netherlands).

Procedure and stimuli

General procedure

Participants were tested in a three-walled canvas test booth placed in a sound-attenuated room. Each infant was seated on the parent's lap approximately 1.35 m from a 17-inch computer screen (Philips LCD 150P4). 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. Parents wore headphones (Echelon Telex), through which music was played in order to prevent them from hearing the stimuli and (potentially) influencing their children's behavior. The experiment was monitored and recorded through a video camera that was placed underneath the TV screen. Looking time was tracked by pressing a button box for looking and looking away. Looking time was taken to reflect listening time (Aslin, 2007). The button box was connected to an Asus P4PE computer. An experiment control application, Zep (Veenker, 2008), was used for presentations of the auditory and visual stimuli and for the data registration. Trials were initiated with a button press and were ended when either the infant looked away for 2 s or the maximum trial length was reached.

Prior to testing, parents provided written consent for participation and the experimental procedure was explained to caregivers without telling them which of the conditions (native or non-native) was presented first; see de Klerk et al., 2019 for further instructions to the caregiver. The aim was to test infants on both contrasts (native and non-native) within one session, and the order was counterbalanced between infants.

Table 1
Numbers of FR participants and mean ages per age group for the native and non-native contrasts.

Age group (months)	Age of FR infants (days)	FR infants tested (n)	FR total infants included (n female)	Of which Both Contrasts (n female)	Only Native Contrast (n female)	Only Non-Native Contrast (n female)	Total included native contrast (n female)	Total included non-native contrast (n female)
6	194 (8.7)	38	33 (14)	26 (11)	2 (1)	5 (2)	28 (12)	31 (13)
8	258 (7.8)	38	32 (15)	19 (6)	5 (3)	8 (6)	24 (9)	27 (12)
10	316 (8.2)	41	33 (19)	20 (10)	7 (7)	6 (2)	27 (17)	26 (12)
Total		117	98 (48)	65 (27)	14 (11)	19 (10)	79 (38)	84 (37)

Note. The numbers in columns 3–7 present the number of infants that were included. The last two columns (8 and 9) represent the datasets (or records) of these infants that were included.

Table 2
Numbers of no-FR participants and mean ages per age group for the native and non-native contrasts.

Age group (months)	Age of no-FR infants (days)	Tested on native contrast only (*n* female)	Tested on non-native contrast only (*n* female)	Tested on both contrasts (*n* female)	Total included native contrast (*n* female)	Total included non-native contrast (*n* female)
6	203 (8.4)	2 (0)	5 (2)	26 (10)	28 (10)	31 (12)
8	259 (6.5)	5 (3)	8 (6)	19 (11)	24 (14)	27 (17)
10	320 (12.9)	7 (4)	6 (2)	20 (10)	27 (14)	26 (12)
Total		14 (7)	19 (10)	65 (31)	79 (38)	84 (41)

Note. These infants are a subsample of the sample presented in de Klerk et al., 2019. no-FR, no family risk of dyslexia.

Table 3
Background information of participants.

Measure	no-FR *M* (*SD*)	FR *M* (*SD*)	Mann–Whitney test (two-sided)
Education level of father	5.58 (0.50)	5.13 (0.90)	$U = 3040.00$, $z = -3.44$, $p = .001$
Education level of mother	5.66 (0.50)	5.48 (0.70)	$U = 3647.00$, $z = -1.77$, $p = .077$
Number of siblings	0.33 (0.50)	0.28 (0.50)	$U = 4000.50$, $z = -0.66$, $p = .509$
Birth rank	1.31 (0.50)	1.27 (0.50)	$U = 4028.00$, $z = -0.57$, $p = .570$

Note. Educational level was measured on a scale from 1 (primary school) to 6 (university/Ph.D.). no-FR, no family risk of dyslexia; FR, at family risk of dyslexia.

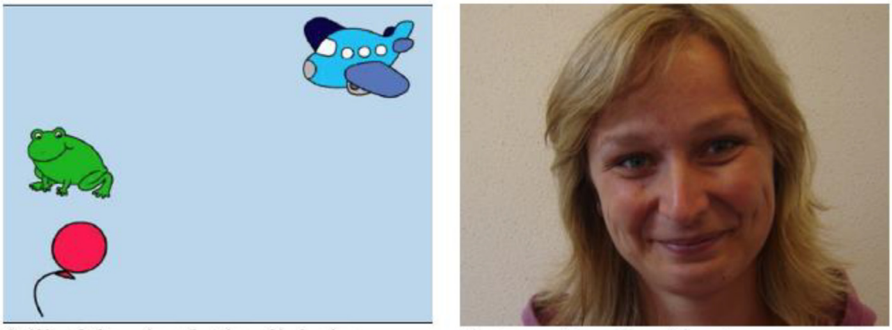
Similar to the study of Houston et al. (2007), the experimental setup consisted of a habituation phase in which infants were habituated to one of the vowels of the pair (e.g., /aː/ in /faːp/), a test phase in which looking times to non-alternating (habituation) vowel pairs (e.g., /faːp/–/faːp/) were compared with those to alternating vowel pairs, that is, a pair consisting of a trained vowel and a contrasting untrained vowel (e.g., /faːp/–/feːp/). The experiment began and ended with a pre- and posttest to measure participants' attentiveness. Each of these phases included both auditory and visual stimuli. During habituation, we used tokens from four different female speakers. Speaker variability has been argued to enhance generalization of abstract features in the process of developing phonetic categories (e.g., Lively, Logan, & Pisoni, 1993; Rost & McMurray, 2009).

Stimuli

Visual and auditory stimuli pre- and posttest. During the pre- and posttest, infants were presented with both auditory stimuli (beep sounds, 330 Hz, duration 250 ms, interstimulus interval [ISI] of 1000 ms) and visual stimuli. Auditory stimuli were played at ~65 dB. The visual stimuli were three cartoon pictures displayed for 2 s on a light blue screen. The three pictures were drawn randomly out of a set of 25 pictures. These pictures could appear in nine different spots within an invisible 3 × 3 grid (see top left picture in Fig. 2.) After 2 s, a series of three new pictures appeared at different locations. Pictures were presented in a pseudorandomized order.

Visual and auditory stimuli habituation and test. During the habituation and test phases, pictures of six smiling female faces were used (see an example in Fig. 2, top right picture). In each block of four trials, four pictures with different female faces were used, one picture per trial. Pictures were presented in a pseudorandomized order. Between habituation trials, a visual attention getter was displayed: a movie of a cute laughing baby (see Fig. 2, bottom left picture). In between test trials, a movie of a toddler going down a slide was used as an attention getter (see Fig. 2, bottom right picture).

Auditory stimuli were the Dutch vowels /aː/ and /eː/ for the native contrast and the English /ɛ/ and /æ/ for the non-native contrast. Vowels were embedded in consonant–vowel–consonant (CVC) syllables: /faːp/, /feːp/, /sæn/, and /sɛn/. Recordings of the Dutch pseudowords /faːp/ and /feːp/ were made of four female Dutch speakers aged 25 to 35 years. They all spoke Standard Dutch and came



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

Fig. 2. Visual stimuli presented during the pre- and posttest, habituation, and test phases. Top left: An example of the visual stimuli during pre- and posttest. Top right: An example of a female face used during habitation and test trials. Bottom left: A still of the attention getter between habituation trials. Bottom right: A still of the attention getter between test trials. (This figure also appeared in de Klerk et al., 2019.)

from the Randstad area, a mostly urban area in central-western Netherlands. They were asked to read out loud a list of 52 words containing the target pseudowords as well as monosyllabic Dutch real words with the same vowels (e.g., *gaap*–yawn, *feest*–party). Recordings of the English pseudowords were recorded of four female native English speakers aged 25 to 35 years. They came from different regions of the United Kingdom: southeast London, Belfast, Preston (Lancashire), and Manchester. The pseudowords /sɛn/ and /sæn/ were read out loud from a list of 52 words containing the target words and real words (e.g., *have* and *pet*).

Each speaker produced four tokens of each target pseudoword (e.g., /faːp/ and /feːp/). From all four speakers, one token of each target pseudoword per contrast was selected. In addition, from one speaker, a second token per target word was selected because this was necessary for the test phase (see “Procedure” section). 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 Fig. 3) was used during the test phase (see “Procedure” section). All auditory stimuli were played at ~65 dB(A). Tokens selected were those that were most child-friendly in prosody and speech affect (see de Klerk et al., 2019 for more details on acoustic properties). All auditory stimuli were recorded in a sound-attenuated booth of the phonetics lab at Utrecht University using a Sennheiser microphone (ME-64) and a digital audio tape recorder (Tascam DA-40).

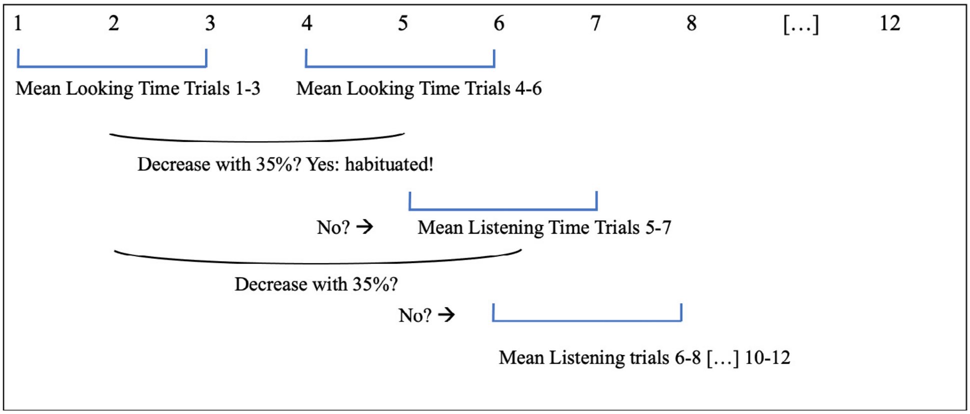


Fig. 3. Visual depiction of the assessment of the (65%) habituation criterion. (This figure also appeared in de Klerk et al., 2019.)

Procedure

Pre- and posttest. The pre- and posttest had a fixed duration of ∼24 s. The purpose of the pre- and posttest was to measure general attentiveness. Infants were excluded when total looking time of the posttest decreased by at least 50% compared with the total looking time of the pretest (n = 1; see "Participants" section above).

Habituation phase. The habituation phase consisted of a maximum of 12 trials, with a maximum number of 30 tokens (ISI of 1 s) per trial, resulting in a total duration of approximately 48 s per trial. Participants were habituated on a repetition of one of the stimulus types (e.g., either /faːp/ or /feːp/ in the native condition and either /sɛn/ or /sæn/ in the non-native condition) with tokens from four female speakers. Within one trial, one token of one speaker was used. In each block of four trials, participants heard all four voices in randomized order within the blocks. Infants were considered to be habituated when they passed the habituation criterion set at 65%: the mean of Trials 1 to 3 was compared with the mean of Trials 4 to 6. If looking time had not decreased by 35%, the mean of the first three trials was compared with the mean looking time of Trials 5 to 7, then Trials 6 to 8, and so on up to Trials 10 to 12 (see Fig. 3).

Test phase. The test phase had a fixed number of 12 trials, with a maximum number of 30 tokens per trial (ISI of 1 s), resulting in a maximum total duration of approximately 48 s per trial. Test trials consisted either of alternating pseudoword pairs (e.g., native /feːp/–/faːp/) or non-alternating pairs (e.g., /faːp/–/faːp/) (see Fig. 4). The alternating and non-alternating trials were presented in a semifixed order; the first trial could be either alternating or non-alternating, which was counterbalanced. The second trial was non-alternating if the first trial was alternating and was alternating if the first trial was non-alternating. The three subsequent alternating trials occurred at positions 5, 8, and 12. The other trials were non-alternating (see Fig. 4). 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., /faːp/–/faːp/, *faap–faap*) had both a new token (*faap*-2 from Speaker 1) and a familiar token (*faap*-1 from Speaker 1); just like in the alternating trials, a new token (*feep*-1 from Speaker 1) and a familiar token (*faap*-1 from Speaker 1) were used (see Fig. 4). The exact same procedure was applied for the non-native contrast.

Offline coding

A random subset (44% of the entire set) of the video recordings was recoded frame by frame (frame duration was 30 ms) using PsyCode software (http://psy.ck.sissa.it/PsyCode/PsyCode.html) by two

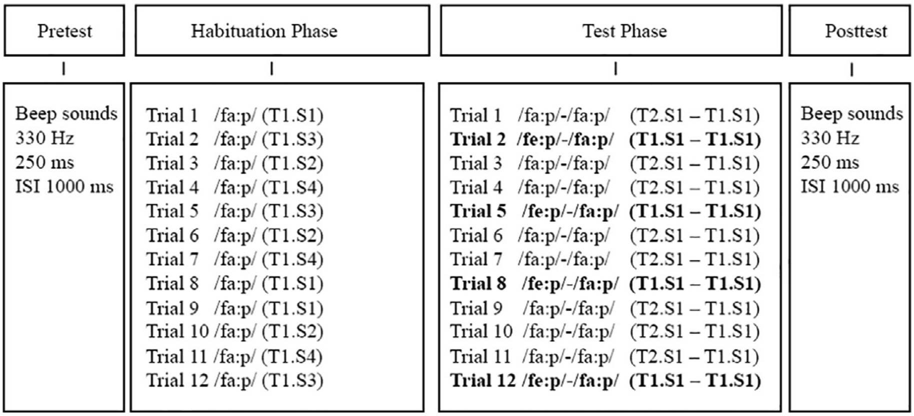


Fig. 4. Schematic overview of the experimental procedure with reference to the auditory stimuli only. In this example, the contrast is native and the first test trial is non-alternating; consequently, the second test trial is alternating. The remaining three alternating trials have a fixed number, namely the 5th, 8th, and 12th trials. Alternating trials are shown in bold. ISI, interstimulus interval; T, token; S, speaker. (This figure also appeared in de Klerk et al., 2019.)

trained coders who were naive regarding the design and purpose of the experiment. The results of the raw and recoded data correlated strongly, $r(105) = .99$, $p < .001$. We used the online coding data for analyses.

Data analysis

Frequentist analyses

To assess whether total looking time and number of trials needed to habituate change as a function of age and/or group (FR or no-FR infants), univariate analyses of variance (ANOVAs) and Kruskal–Wallis nonparametric tests were conducted. Random effects modeling (SPSS Version 23) was used to answer the questions of (a) whether there was an effect of trial type (alternating [e.g., /faːp/–/feːp/] or non-alternating [e.g., /faːp/–/faːp/] trials), (b) whether there were differences between the age groups, and (c) whether there were differences between the groups (no-FR and FR). The overall fit of the model was tested with a chi-square likelihood ratio test. Seven trials were not included for data analysis; one infant missed the last two trials because the experiment was terminated, and five trials were excluded because (three different) infants were suspected of gazing and not looking at the screen. The missing data were not considered problematic because (a) very few trials were missing and (b) parameters can be estimated accurately with missing data using mixed modeling (Field, 2013). For all the frequentist analyses reported in this study, the alpha level was .05.

Bayesian analysis

In our Bayesian hierarchical regression model, we modeled the individual infant data in three age groups (6, 8, and 10 months) per group (FR and no-FR infants) and contrast (native and non-native), as we did in our previous study (de Klerk, Veen et al., 2019). In that study, we presented all details, notably priors, estimation and convergence, and posterior predictive checking, and we conducted a sensitivity analysis (see https://osf.io/xyh3g/). We used a regression model with an AR1 error structure, with $\log_{10}$ transformed looking times as outcomes and condition (alternating or non-alternating trials) as predictor. For all groups, we obtained both group and individual estimates for the intercept (looking time alternating trials) and condition (difference in looking time between alternating and non-alternating trials).

Results

Group analyses: Random effect modeling

Data screening

Habituation phase. The mean of the total looking times to habituation trials as well as the number of trials required for habituation were assessed across ages. The looking time distributions were positively skewed. Log transformation ($\log_{10}$) resulted in a distribution that approached a normal distribution (skewness = .026, *SE* = .039, kurtosis = .488, *SE* = .078). The mean number of trials needed to habituate did not approach a normal distribution after log transformation. Therefore, nonparametric tests were conducted on this measure.

Test phase. The raw looking times to alternating and non-alternating trials were not normally distributed; for this reason, a $\log_{10}$ transformation was performed. After this transformation, the skewness (.096, *SE* = .039) and kurtosis (.256, *SE* = .078) values were acceptable.

The effect of contrast on discrimination

Habituation phase. Mean looking times required for habituation are reported in Fig. 5. Analyses yielded a significant main effect of age, $F(2, 313) = 4.51$, $p = .012$. Post hoc analyses showed that the 10-month-olds had overall shorter looking times than the 6- and 8-month-olds. No other main effects and interactions between contrast and age or group were found. The total looking times to habituation trials did not differ between contrasts or between FR and no-FR infants. The mean number of trials needed to habituate are presented in Table 4. < T4 > A Kruskal–Wallis test revealed no differences on these mean numbers of trials between no-FR and FR infants in the native contrast, $H(1) = .23$, $p = .637$, or in the non-native contrast, $H(1) = .24$, $p = .626$. Because there were no significant differences between the no-FR and FR infants regarding habituation for both contrasts, we do not discuss habituation separately per contrast.

Test phase. We first investigated the effect of contrast (native or non-native) to find out whether trajectories differed between contrasts. Significant interactions of contrast with trial type and/or age would lead us to analyze the results per contrast separately. Looking times per trial type (alternating or non-alternating trials) are presented in Fig. 6. 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 or non-alternating trials), contrast (native or non-native), and age (6,

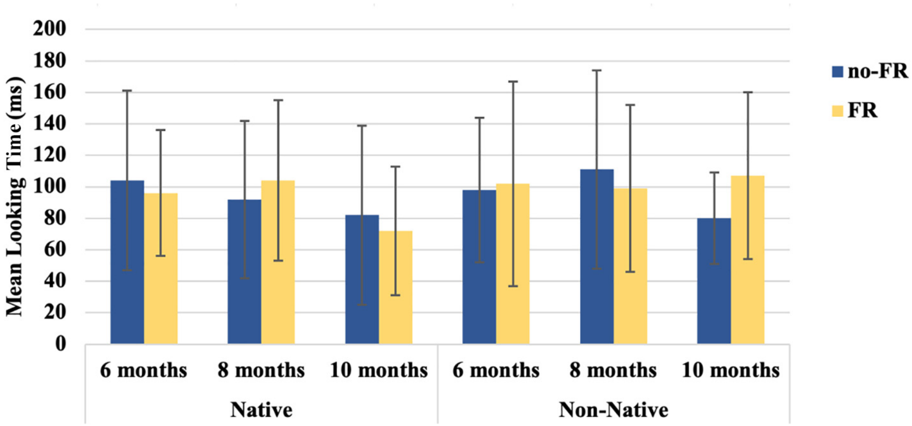


Fig. 5. Mean looking times to habituation trials per contrast and group. Error bars represent standard deviations. no-FR, at no family risk; FR, at family risk.

Table 4
Mean numbers of trials needed to habituate per contrast and group.

Age group (months)	Contrast	Group	
		no-FR	FR
	Native	*M* (*SD*)	*M* (*SD*)
6		6.6 (1.3)	6.5 (1.2)
8		6.7 (1.5)	7.3 (1.8)
10		6.6 (1.1)	6.8 (1.9)
Total		6.6 (1.3)	6.9 (1.6)
	Non-native		
6		6.7 (1.5)	7.2 (1.8)
8		7.7 (2.0)	7.1 (1.7)
10		7.1 (1.7)	7.5 (2.0)
Total		7.1 (1.8)	7.3 (1.8)

Note. no-FR, no family risk of dyslexia; FR, at family risk of dyslexia.

8, or 10 months). The model that best fitted the data included the fixed factors trial type (alternating or non-alternating trials), *F*(1, 2729) = 88.26, *p* < .001, contrast (native or non-native), *F*(1, 338) = 1.16, *p* = .282, age, *F*(1, 338) = 3.94, *p* = .020, Trial Type * Contrast, *F*(1, 2729) = 8.24, *p* = .004, and Trial Type * Contrast * Age, *F*(6, 906) = 2.78, *p* = .011. The two-way and three-way interactions show that the effect of trial type on looking time varied across contrasts and ages. Therefore, separate analyses for each contrast are presented in the next sections. The main effect of trial type indicates that the infants listened longer to alternating trials than to non-alternating trials. Looking times decreased with age, as indicated by the main effect of age. No main effect of contrast was found, indicating that the overall looking times were not significantly different for the two contrasts.

The native contrast

Test phase. Looking times are reported in Table 5. 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 or non-alternating trials), age (6, 8, or 10 months), group (no-FR or FR infants), and habituation stimulus (/faːp/ or /feːp/). A two-way interaction (Trial Type * Group) or a three-way interaction (Trial Type * Age * Group) would show that groups (no-FR and FR) responded

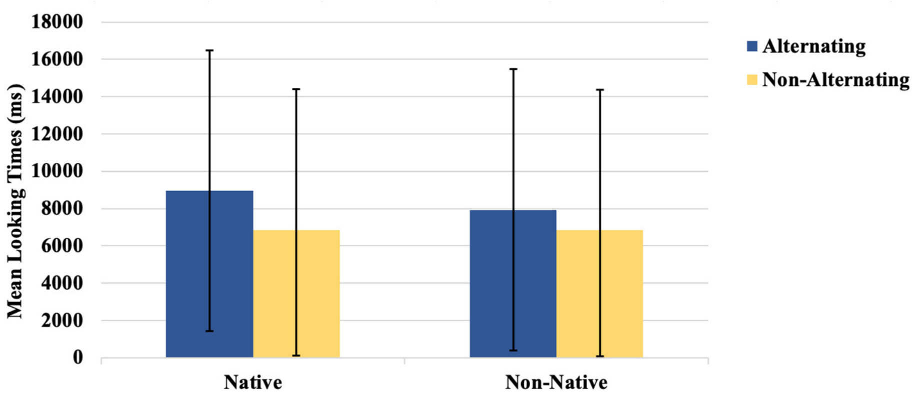


Fig. 6. Mean looking times to alternating and non-alternating trials for the native and non-native contrasts. Error bars represent standard deviations.

Table 5
Looking times to alternating and non-alternating trials of the native contrast per group.

Group	Age	Alternating trials		Non-alternating trials		Statistics			Participants	Infants with longer looking times to alternating trials[a]	
	(months)	*M*	(*SD*)	*M*	(*SD*)	*F*	*p*	Cohen's *d*	(*N*)	*n*	%
no-FR	6	9.9	(5.2)	8.0	(3.1)	9.04	.003	0.49	28	17	60
	8	9.2	(6.0)	6.4	(3.0)	13.32	<.001	0.66	24	15	62
	10	7.6	(3.7)	5.4	(2.4)	21.33	<.001	0.76	27	20	74
	All	8.9	(5.1)	6.6	(3.0)	42.07	<.001	0.60	79	52	66
FR	6	9.7	(6.0)	7.5	(2.8)	4.66	.031	0.53	28	19	68
	8	8.2	(4.9)	6.4	(3.1)	5.54	.019	0.48	24	18	75
	10	8.9	(4.5)	7.3	(4.4)	14.13	<.001	0.36	27	17	63
	All	9.0	(5.1)	7.1	(3.5)	23.47	<.001	0.46	79	54	68

Note. Looking times are given in seconds. no-FR, no family risk of dyslexia; FR, at family risk of dyslexia.

[a] Number of infants who had on average longer looking times to alternating trials than to non-alternating trials.

differently to the alternating and non-alternating trial types and would indicate differences in discrimination performance.

The model that best fitted the data included the fixed factors trial type (alternating or non-alternating trials) and age (6, 8, or 10 months). This model yielded significant effects of (a) trial type on looking time, *F*(1, 1346) = 71.63, $p < .001$, indicating that infants listened longer to alternating trials than to non-alternating trials, and (b) age, *F*(2, 158) = 5.58, $p = .005$, indicating that overall looking times decreased as age increased. As can be seen in Table 5, within the no-FR and FR groups, all age groups discriminated the non-native vowel contrast. However, the effect sizes in Table 5 suggest that there were differences between the groups. The effect size (Cohen's *d*) per age group for the no-FR infants increased from a moderate value (.49) to a large value (.76). This was not the case for the FR infants, whose effect size dropped from a moderate effect size of .53 at 6 months to a moderate to small effect size of .36 at 10 months. Large variations in looking times resulted in smaller effect sizes, implying a less robust effect of trial type.

In sum, both groups of infants were able to generalize over speaker variations during habituation and responded to those acoustic features that differentiate between Dutch /aː/ and /eː/ regardless of whether the habituation stimulus was /faːp/ or /feːp/. However, whereas for the no-FR infants there was an increase in the effect size of the mean difference between alternating and non-alternating trials across age, this was not seen for the FR infants due to the 10-month-olds who showed more variance between infants.

The non-native contrast

Test phase. Table 6 displays the results of the test phase. 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 or non-alternating trials), age (6, 8, or 10 months), group (no-FR or FR infants), and habituation stimulus (/faːp/ or /feːp/).

The model that best fitted the data included the fixed factors trial type (alternating or non-alternating trials), age (6, 8, or 10 months), group (no-FR or FR infants), and habituation stimulus (/sæn/ or /sɛn/). Infants looked longer to alternating trials than to non-alternating trials, *F*(1, 1321) = 12.63, $p < .001$. The significant Trial Type * Age * Group interaction, *F*(7, 545) = 3.69, $p = .001$, was explored by Bonferroni-adjusted pairwise comparisons. No-FR 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 6). FR infants showed no evidence of discrimination in any of the age groups. The Trial Type * Habituation Stimulus * Group interaction was also significant, *F*(3, 435) = 5.59, $p = .001$. Post hoc analyses showed that the no-FR infants discriminated the contrast regardless of habituation stimulus. FR infants did not discriminate the contrast, also regardless of habituation stimulus. The fixed factor habituation stimulus yielded no main effect, *F*(1, 158) = 1.33,

Table 6
Looking times to alternating and non-alternating trials of both groups for the non-native contrast.

Group	Age	Alternating trials		Non-alternating trials		Statistics			Participants	Infants with longer looking times to alternating trials[a]	
	(months)	*M*	(*SD*)	*M*	(*SD*)	*F*	*p*	Cohen's *d*	(*N*)	*n*	%
no-FR	6	8.6	(4.9)	6.4	(3.6)	11.37	.001	0.39	31	25	80
	8	7.0	(2.7)	7.1	(3.1)	0.28	.599	−0.03	27	14	51
	10	8.4	(4.2)	6.0	(2.1)	20.82	<.001	0.44	26	22	85
	All	8.1	(4.1)	6.4	(3.0)			0.45	84	61	73
FR	6	8.5	(6.1)	7.6	(2.3)	1.40	.237	0.23	31	17	55
	8	7.1	(3.5)	7.3	(2.7)	0.63	.427	−0.07	27	11	41
	10	8.4	(5.2)	7.7	(5.1)	3.01	.081	0.14	26	18	69
	All	8.0	(5.1)	7.6	(3.6)	2.26	.133	0.10	84	46	55

Note. Looking times are given in seconds. no-FR, no family risk of dyslexia; FR, at family risk of dyslexia.
[a] Number of infants who had on average longer looking times to alternating trials than to non-alternating trials.

$p = .250$, and no significant interaction between trial type and habituation stimulus was found, $F(1, 1251) = 0.09$, $p = .767$, or among trial type, habituation stimulus, and age, $F(5, 356) = 0.67$, $p = .647$. There was no effect of age, $F(2, 157) = 0.07$, $p = .935$, meaning that no evidence was found for a difference in overall looking times among age groups. The main effect of group was not significant, $F(2, 157) = 0.36$, $p = .551$.

Unlike the results of the native contrast, performances of the no-FR and FR groups clearly differ. The 6- and 10-month-old no-FR infants showed evidence of discrimination of the non-native vowel, whereas the 8-month-olds did not. The FR infants did not show evidence of non-native discrimination at any of the ages.

Individual analyses: Bayesian hierarchical modeling analyses

Findings of the Bayesian hierarchical regression model are presented in Table 7. The parameter of interest was trial type (alternating or non-alternating trials) because this allowed us to establish whether the looking times differed between the alternating and non-alternating conditions for the individual infants. The aim was to classify infants as discriminators or nondiscriminators. Using a

Table 7
Numbers and percentages of infants whose 95% CIs did not include the value zero.

Bayesian Hierarchical modeling				
			Number (and %) of infants whose 95% CIs did not include zero	
Contrast	Age group	Total number of participants	Group	
	(months)		no-FR	FR
Native	6	55	3/28 (11%)	1/27[a] (4%)
	8	48	4/24 (17%)	16/24 (67%)
	10	54	15/27 (56%)	6/27 (22%)
	Subtotal	158	20/79 (25%)	23/79 (29%)
Non-native	6	61	19/30[a] (63%)	0/31 (0%)
	8	53	0/27 (0%)	0/26[a] (0%)
	10	51	21/25[a] (84%)	0/26 (0%)
	Subtotal	168	41/84 (49%)	0/84 (0%)
	Total	326	61/163 (37%)	0/163 (0%)

Note. Data of 4 participants were not included in the Bayesian analysis due to missing data. The superscript letter "a" in the no-FR (no family risk of dyslexia) and FR (at family risk of dyslexia) columns indicates age groups for which data were missing. CI, credibility interval.

frequentist approach, this would mean that an individual is classified as being able to discriminate a contrast if the mean difference between alternating and non-alternating trials differs significantly from zero. Here we followed a similar criterion. We checked which infants had the value zero in their 95% credibility interval (CI) for the trial type parameter; inclusion of value zero means that there is no evidence for a difference in looking times between the two trial types (alternating and non-alternating). See Table 7 for the percentages of infants for whom the 95% CI did not cross zero, and see Appendices C to F for the group and individual estimates and the variance per group per contrast.

Native vowel contrast

Of the no-FR infants, 11% of the 6-month-olds, 17% of the 8-month-olds, and 56% of the 10-month-olds did not have zero in their 95% CIs (see Table 7). Hence, these values are taken to reflect the percentages of children who discriminated between the native vowels. Fig. C1 in Appendix C shows the estimated medians and CIs. The data indicate that the individual 95% CIs for the 8-month-olds show larger uncertainty than the individual 95% CIs for the other age groups. Fig. C2 in Appendix C shows that the variance estimates are larger for the 8-month-olds than for the other age groups; the 8-month-olds differ more from one another than the 6- and 10-month-olds. Larger variance at the group level influences the individual estimates because these become more uncertain. Hence, fewer infants can be classified as discriminators. Figs. C1 and C2 show that the estimated effect of condition in the 10-month-olds is comparable to that in the 8-month-olds. The higher percentage of infants who do not have zero in their CIs is due to the smaller variance in the 10-month-olds as a group. They resemble one another more than do the 8-month-olds; therefore, we are more confident about their estimated condition effects at the individual level. Hence, this might indicate that discrimination of the native /aː/–/eː/ contrast becomes more robust with maturation.

The percentages of the FR infants who do not have zero in their CIs, and thus are considered to discriminate the native contrast, are 4% for the 6-month-olds, 67% for the 8-month-olds, and 22% for the 10-month-olds (see Table 7). Figs. D1 and D2 in Appendix D show that the group estimates of the condition effect are similar for the 8- and 10-month-olds. The 10-month-old FR infants show more variance in the group estimates of the condition effect (Fig. D2). Hence, they show larger uncertainty in their individual 95% CIs, similarly to the no-FR 8-month-olds. The finding that, compared with the 8-month-old FR infants, so few 10-month-old FR infants can be classified as discriminators is due to large uncertainty in the individual estimates of the 95% CI. As in the 10-month-olds, very few 6-month-olds have CIs that do not include the value zero. The group variances of the 6- and 8-month-olds (right panel of Fig. D2) are comparable. So, variance at the group level cannot explain the difference in the percentages of infants (6-month-olds vs. 8-month-olds) who do not include the value zero in their CIs. However, the estimate of the condition effect and the CI of this effect are closer to zero (left panel of Fig. D2) for the 6-month-olds. In this hierarchical model, the estimated effect of condition at the group level functions as a prior for the individual condition effect estimates, and individual estimates are pooled toward these group estimates. The literature has shown that incorporating group structures into the analyses leads to fewer mistakes for the individual parameters that are estimated in terms of the magnitude and sign (direction) of the effects. This issue is addressed more elaborately in the literature on Type S and Type M errors, mostly by Gelman and Tuerlinckx (2000). Thus, keeping the group level variance equal, a smaller group-level estimate for the condition effect will pool the individuals toward smaller individual effects, as can be seen in the 6-month-olds. At the group level, we found evidence for discrimination and homogeneity of the group. Because the group estimate and CI are closer to zero, the individual estimates are as well.

To summarize, the no-FR infants can discriminate the native contrast. The data suggest an enhancement effect; the percentage of infants who discriminate the contrast at 6 months of age is low, but this increases with age. The FR infants show a different developmental pattern. The 6-month-olds show evidence of discrimination at the group level, but very few individuals can be classified as discriminators. The 8- and 10-month-olds are able to discriminate the native contrast at the group level. At the individual level, however, too much uncertainty remains to classify many of the 10-month-olds as discriminators.

Non-native vowel contrast

The percentages of no-FR infants who do not have zero in their 95% CIs are 63% for the 6-month-olds, 0% for the 8-month-olds, and 84% for the 10-month-olds. In Figs. E1 and E2 in Appendix E, it can be seen that the low percentage of the 8-month-olds cannot be attributed to larger uncertainty of the CIs or to the larger variation of the group estimates. The individual estimates of the median condition effect are close to zero, and the 95% CIs convincingly cross zero.

The FR infants show different results; for all age groups, none of the FR infants discriminated between the non-native vowels given that they all have zero in their 95% CIs. These results cannot be explained by larger variability or uncertainty; the individual estimates of the means are all at or close to zero (see Figs. F1 and F2 in Appendix F). However, although the percentage of the 10-month-old infants who do not have zero in their CIs is 0%, the group estimate for the condition effect is not close to zero, which indicates some effect of condition at the group level. In addition, not all the individual mean estimates are close to zero (see right panel of Fig. F1). Hence, some CIs barely cross the zero, as was seen in the 6-month-old FR infants in the native condition. The 6- and 8-month-old FR infants did not discriminate the non-native contrast at the group level or at the individual level. At 10 months of age, we found discrimination at the group level, but the uncertainty at the individual level means that we cannot classify any individual as a discriminator.

Together, the cross-sectional data of the no-FR infants suggest a U-shaped pattern of discrimination of the non-native contrast (discrimination at 6 months, not at 8 months, but again at 10 months), similar to our previous findings (de Klerk et al., 2019; de Klerk, Veen et al., 2019). The 6- and 8-month-old FR infants did not show evidence of discrimination at the individual level or at the group level. However, as was seen with the 6-month-olds in the native condition, the 10-month-olds did show some evidence of discrimination at the group level but not at the individual level.

Discussion

The aim of this study was twofold. The first was to evaluate whether a similar pattern of perceptual attunement would be attested for children with and without a family risk of dyslexia. The second was to assess whether this pattern was reflected in group findings as well as individual-based analysis. To look into these questions, discrimination of native and non-native phonemes in 6-, 8-, and 10-month-old infants was studied. Perceptual attunement would be attested if (a) the native contrast (/aː/–/eː/) was discriminated at all ages or when discrimination improved with age and (b) discrimination performance of the non-native contrast (English /ɛ/–/æ/) declined with age. In light of the proposed speech perception deficit in children with (an FR of) dyslexia (e.g., Molfese, 2000; Richardson et al., 2003; Werker & Tees, 1987), it was expected that the FR infants would show evidence of discriminating the salient native contrast but that this discrimination would show a slower improvement than in the no-FR infants. With respect to the subtle non-native contrast, it could be the case that the FR group would not discriminate this contrast at any time point (e.g., Richardson, et al., 2003; van Leeuwen et al., 2006) or that there would be continued discrimination as opposed to a decrease in the no-FR group (Noordenbos et al., 2012).

There was no evidence for a difference between the no-FR and FR infants on the native speech contrast. The (frequentist) group findings of the native contrast showed that both the no-FR and FR infants discriminated salient Dutch /aː/ from /eː/. However, the effect sizes show subtle differences between the two groups; whereas there was an increase in effect sizes over age for no-FR infants between alternating and non-alternating trials at the group level, this was not found for the FR infants. Hence, the variability reduces with age in no-FR infants. This could be indicative of an increasingly robust discrimination performance with age/maturation, aligning with theories of enhancement of native speech perception (Kuhl et al., 2008; Tsuji & Cristia, 2014). The FR group does not show this increase, which may imply that there is a subtle delay. A longitudinal study extending to older age groups could be used to investigate whether an enhancement effect does surface for the FR group at a later age.

The Bayesian individual outcomes showed that the percentage of no-FR infants who discriminate the native contrast at 6 months of age is low and increases with age. The observation that the 8-month-olds, compared with the 10-month-olds, showed a relatively low percentage of discriminators is explained by the larger variance in the group estimates. The pattern of the FR infants at the individual level is different. The 6-month-olds showed (very) weak evidence of (native) discrimination, and a relatively low percentage of 10-month-olds could be classified as discriminators compared with the 8-month-olds. The explanation is the same here as it is for the 8-month-old no-FR infants, namely that the variance for the 10-month-old FR infants is larger than that for the 8-month-olds. Larger variance at the group level influences the individual estimates as these become more uncertain. Hence, fewer infants can be classified as discriminators. The 10-month-old FR infants as a group behave less coherently given that they differ more from one another, matching the findings from the frequentist analysis. This could be indicative of a subtle delay in speech perception development.

The outcomes of the non-native contrast showed a different picture, with pronounced differences between the no-FR and FR groups. Both the frequentist group analysis and the Bayesian individual analysis suggest a pattern of U-shaped development for the no-FR group but not for the FR group. The 6- and 10-month-old no-FR infants showed evidence of discrimination, whereas the 8-month-olds did not. These findings seem to confirm the findings of perceptual attunement between 6 and 8 months of age (Kuhl et al., 1992; Polka & Werker, 1994). In a previous study (de Klerk et al., 2019), we proposed that the improved performance of infants at 10 months is due to their being better equipped than 6- and 8-month-olds to make use of the speaker variation presented during the habituation phase.

For the FR infants, the frequentist group analysis showed no evidence of discrimination in any of the age groups. This was mirrored in Bayesian hierarchical modeling, where none of the FR infants could be classified as discriminators at any age. These findings are difficult to relate to the classical view on perceptual attunement. However, the literature on perceptual attunement has shown that salience influences the ability to discriminate, with more subtle and less salient contrasts needing to be acquired through language exposure (e.g., Liu & Kager, 2015; Narayan, Werker, & Beddor, 2010). Hence, the finding that even the 6-month-old FR infants did not show evidence of discrimination could be due to a lack of initial sensitivity to subtle contrasts and, because infants are not exposed to this non-native contrast, their discrimination performance does not improve. The current findings indicate that it might be important to investigate the developmental trajectory of discrimination performance of subtle *native* contrasts. Based on the data presented in this study, a delay in discrimination performance is expected.

FR infants showed evidence of discriminating the native contrast but not the non-native contrast. Hence, the data of the current study support the notion that FR infants have a (subtle) speech perception deficit. This is in line with studies that investigated speech perception at an early age (e.g., Richardson et al., 2003) and also with studies that investigated the relation between early speech sound processing and later reading outcomes retrospectively (e.g., Molfese, 2000; van Zuijen, et al., 2013). Outcomes of the current study do not support the hypothesis that infants remain sensitive to irrelevant non-native contrasts (Noordenbos et al., 2012). Furthermore, the data of the current study suggest a subtle delay in the development of speech sound categories. There is evidence that distributional learning plays a critical role in the acquisition of native speech sound categories (Kuhl et al., 1992; Maye, Weiss, & Aslin, 2008). Notably, Maye et al. (2008) showed that the frequency distribution of non-native speech sound tokens (unimodal or bimodal) that differ along an acoustic parameter determines whether 8-month-old infants assign them to one or two classes. We hypothesize that FR infants are less proficient in exploiting such distributional information (see Kerkhoff, de Bree, de Klerk & Wijnen, 2013; Wijnen, 2013). This is a hypothesis we are currently investigating.

There are three findings in the current study that require clarification. The first is the finding that the 8-month-old no-FR infants have a lower percentage of discriminators of the native contrast compared with the 6- and 10-month-olds. This finding, suggestive of a U-shaped development of native vowel discrimination, is due to the larger variance at the group level at 8 months. In line with Werker, Hall, and Fais (2004), we propose that the heterogeneous performance at this age reflects a developmental (reorganizational) shift in vowel perception and that it does not reflect a loss of discrimination. The shift refers to a change in processes and strategies applied during speech perception. Younger infants react to all perceivable phonetic differences and thus discriminate all speech sounds. During the reorganizational phase, they begin to learn to categorize speech sounds in phonemic units (Kuhl, 2004). Because not all infants begin their shift at exactly the same time point, this could explain the heterogeneity of group performance at 8 months of age. Although the reorganization in speech perception has so far been observed in tasks assessing discrimination of non-native speech sounds that are assimilated to native speech sounds (Maurer & Werker, 2014), we submit that our individual-based analysis also captured this pattern for a salient native contrast. Hence, we argue that the U-shaped pattern reveals the underlying process of speech perception, which is in line with other studies that interpret U-shaped findings (e.g., Bjorklund, Miller, Coyle, & Slawinsky, 1997; Kachel, Hardecker, & Bohn, 2021; Pauls, Macha, & Petermann, 2013; Siegler, 2004).

The second finding that requires further consideration is related to this first issue of a "dip" in performance of the 8-month-olds in the native speech contrast. It concerns the finding that no enhancement effect was found for the 10-month-old FR infants in the native contrast. Instead, the analyses at the individual level showed a declining percentage of infants that can be classified as discriminators. We would expect more robust discrimination of the native contrast by the 10-month-old FR infants on the basis of the finding that 8-month-old FR infants are able to discriminate this contrast. The 10-month-olds have had more experience with their native language and the salient native contrast and therefore would show better discrimination. However, if FR infants indeed have a subtle delay in the development of speech sound categories, as our data suggest, the lower percentage of FR infants who discriminate the native contrast at 10 months of age could be an indicator of them being at the reorganizational phase, similar to the 8-month-old no-FR infants. Further research in the underlying processes of this reorganization shift would be welcome.

The third finding that needs clarification is that the no-FR 10-month-olds could discriminate the non-native contrast given that this was not anticipated on the basis of perceptual attunement. This cannot be due to exposure because infants are not exposed to this non-native contrast in real life. We propose that the 10-month-old no-FR infants could discriminate this contrast because they were better able to make use of the speaker variation presented during the habituation phase. Speaker variation stimulates phonetic learning because it demands abstraction of invariant features (Lively et al., 1993), but this effect is likely to become stronger as age increases (see also de Klerk et al., 2019; Rost & McMurray, 2009). The fact that this effect was not seen in the FR infants might indicate that this phonetic distributional information caused by speaker variation is not helpful to the same extent for FR infants at this age as it is for their no-FR peers.

We have shown that individual analysis can be used to infer whether infants in two different groups are discriminators of native and non-native speech sound contrasts. Moreover, we found large differences between the two groups of infants, indicating that the method presented here could be used to study language(-related) development prospectively. Although previous studies have connected infant speech perception data to later language outcomes (e.g., Newman, Ratner, Jusczyk, Jusczyk, & Dow, 2006; Tsao, Liu, & Kuhl, 2004), using Bayesian hierarchical modeling might render even more sensitive results because it is able to produce both group and individual estimates and could be extended in a straightforward manner to address prospective research questions, relating early speech perception to later language outcomes. For example, the results of the individual estimates can be used in a prospective longitudinal design in which language outcomes, such as vocabulary size, mean length of utterance, sentence complexity, and reading outcomes, are predicted by the (amount of) discrimination performance at this early age. Thus, future studies can provide valuable input on the question of whether discrimination skills at an early age are associated with reading problems or are instead a risk factor (endophenotype) for developing dyslexia (Moll, Loff, & Snowling, 2013).

Although the method presented here for identifying discriminators seems to be a fruitful avenue for gaining more understanding of development, group differences, and prospective relationships, some limitations need to be mentioned. In the current approach, individual and group estimates are influencing one another. It is desirable to obtain a sufficiently large sample size to estimate group-level parameters with confidence. Another limitation is that we did not assess discrimination performance longitudinally and relied on cross-sectional data. Testing speech sound discrimination longitudinally would allow us to establish whether the U-shaped patterns are also attested in such a sample and would provide more insight into whether individual classification is as valuable as we take it to be.

To conclude, we hope to have shown that individual analysis in speech discrimination experiments with infants is a promising avenue for further research. There are still some challenges using the Bayesian hierarchical approach, but it provides us with a tool that allows better understanding how speech perception develops at an individual level (preferably longitudinally) as well as looking prospectively at the relationship with other facets of language and literacy development.

Acknowledgments

We thank the parents of our participants for their contributions to our study. We thank Sule Kurtçebe, Tinka Versteegh, Lorijn Zaadnoordijk and Joleen Zuidema for assistance in data collection and Annemarie Kerkhoff for assistance in the design of the study. This study was funded by The Netherlands Organization for Scientific Research (NWO, grant nr. 360-70-270).

Appendix A. Numbers of participants divided by habituation stimulus and contrast order in the native condition

Age group (months)	Group	Participants (*N*)	Habituation stimulus /fa:p/	Native first[a]
6	No-FR	28	13/28	16/28
8		24	9/24	15/24
10		27	20/27	15/27
Total		79	42/79	46/79
6	FR	28	13/28	16/28
8		24	6/24	14/24
10		27	13/27	8/27
Total		79	32/79	38/79

Note. no-FR, no family risk of dyslexia; FR, at family risk of dyslexia.

[a] Number of participants who received the native contrast first during the test session.

Appendix B. Numbers of participants divided by habituation stimulus and the contrast order in the non-native condition

Age group (months)	Group	Participants (*N*)	Habituation stimulus /sæn/	Non-native first[a]
6	no-FR	31	18/31	18/31
8		27	18/27	16/27
10		26	14/26	17/26
Total		84	50/84	51/84
6	FR	31	17/31	17/31
8		27	19/27	15/27
10		26	11/26	24/26
Total		84	47/84	56/84

Note. no-FR, no family risk of dyslexia; FR, at family risk of dyslexia.

[a] Number of participants who received the non-native contrast first during the test session.

Appendix C. Individual and group estimates for the no-FR infants and native contrast

See Figs. C1 and C2.

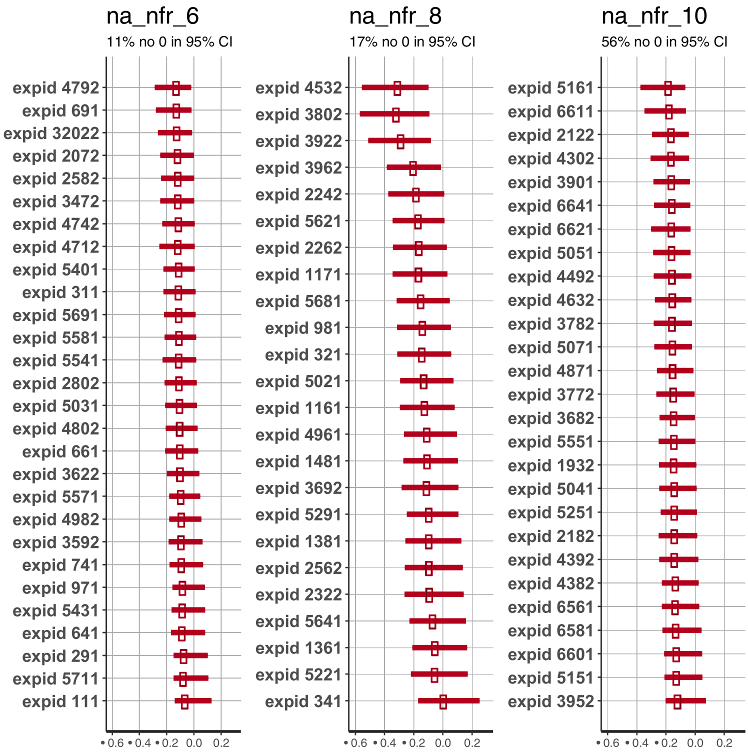


Fig. C1. Results of the hierarchical model for each individual per age group. The rectangles represent the means; the red bars represent the 95% credibility intervals (CIs). no-FR, no family risk of dyslexia.

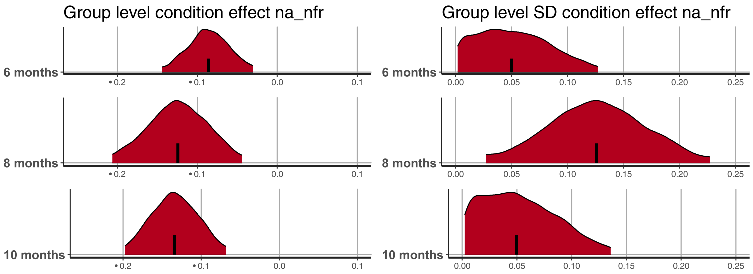


Fig. C2. Group estimates for condition effects and variation per age group. The left panel shows the group estimates for condition effects. The right panel shows the standard deviation (SD) of the condition effect per age group. The densities, presented in red, represent the 95% credibility intervals. no-FR, no family risk of dyslexia.

Appendix D. Individual and group estimates for the FR infants and native contrast

See Figs. D1 and D2.

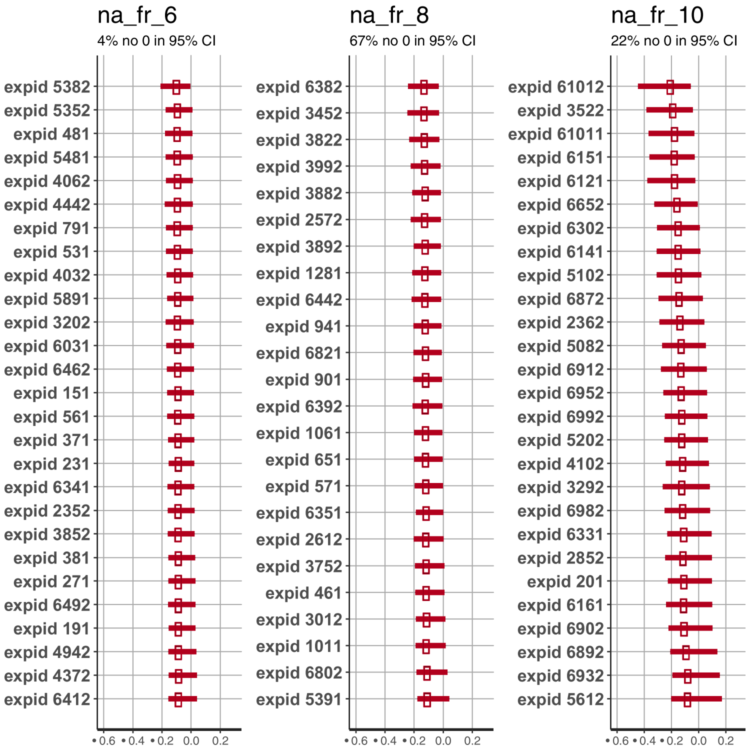


Fig. D1. Results of the hierarchical model for each individual per age group. The rectangles represent the means; the red bars represent the 95% credibility intervals (CIs). FR, at family risk of dyslexia.

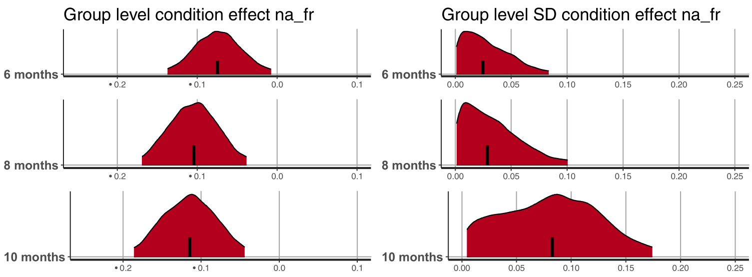


Fig. D2. Group estimates for condition effects and variation per age group. The left panel shows the group estimates for condition effects. The right panel shows the standard deviation (SD) of the condition effect per age group. The densities, presented in red, represent the 95% credibility intervals. FR, at family risk of dyslexia.

Appendix E. Individual and group estimates for the no-FR infants and non-native contrast

See Figs. E1 and E2.

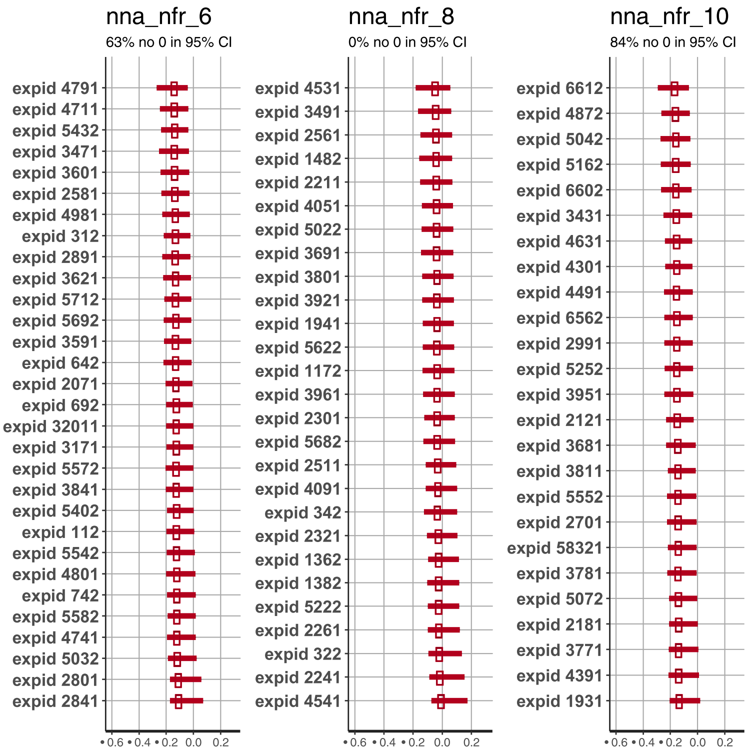


Fig. E1. Results of the hierarchical model for each individual per age group. The rectangles represent the mean; the red bars represent the 95% credibility intervals (CIs). no-FR, no family risk of dyslexia.

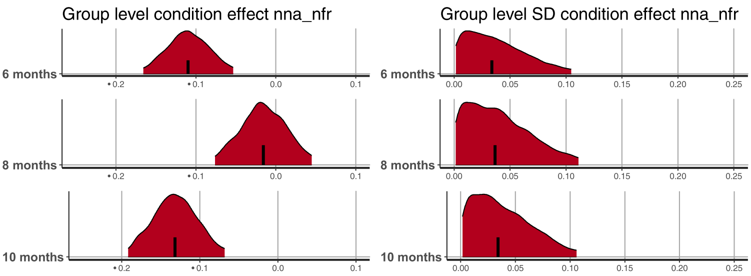


Fig. E2. Group estimates for condition effects and variation per age group. The left panel shows the group estimates for condition effects. The right panel shows the standard deviation (SD) of the condition effect per age group. The densities, presented in red, represent the 95% credibility intervals. no-FR, no family risk of dyslexia.

Appendix F. Individual and group estimates for the FR infants and non-native contrast

See Figs. F1 and F2.

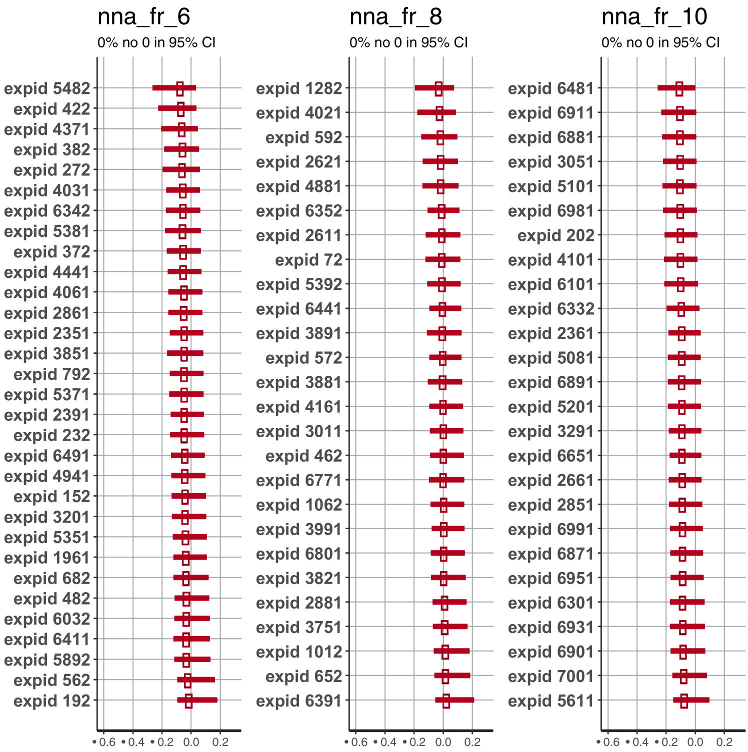


Fig. F1. Results of the hierarchical model for each individual per age group. The rectangles represent the means; the red bars represent the 95% credibility intervals (CIs). FR, at family risk of dyslexia.

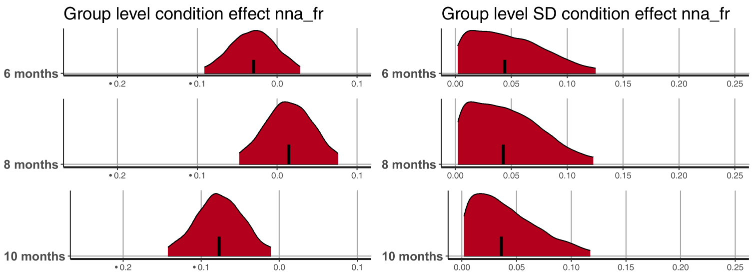


Fig. F2. Group estimates for condition effects and variation per age group. The left panel shows the group estimates for condition effects. The right panel shows the standard deviation (SD) of the condition effect per age group. The densities, presented in red, represent the 95% credibility intervals. FR, at family risk of dyslexia.

References

Aslin, R. N. (2007). What's in a look?. *Developmental Science, 10*, 48–53.

Bjorklund, D. F., Miller, P. H., Coyle, T. R., & Slawinski, J. L. (1997). Instructing children to use memory strategies: Evidence of utilization deficiencies in memory training studies. *Developmental Review, 17*, 411–441.

Blomert, L. (2011). The neural signature of orthographic–phonological binding in successful and failing reading development. *NeuroImage, 57*, 695–703.

Blomert, L., & Mitterer, H. (2004). The fragile nature of the speech– perception deficit in dyslexia: Natural vs. synthetic speech. *Brain and Language, 89*, 21–26.

Boets, B., Ghesquière, P., Van Wieringen, A., & Wouters, J. (2007). Speech perception in preschoolers at family risk for dyslexia: Relations with low-level auditory processing and phonological ability. *Brain and Language, 101*, 19–30.

Brandt, J., & Rosen, J. J. (1980). Auditory phonemic perception in dyslexia: Categorical identification and discrimination of stop consonants. *Brain and Language, 9*, 324–337.

Broersma, M., & Cutler, A. (2011). Competition dynamics of second-language listening. *Quarterly Journal of Experimental Psychology, 64*, 74–95.

Brus, B. T., & Voeten, M. J. M. (1972). *Een-Minuut-Test: Vorm A en B* [One-Minute Test]. Nijmegen, Netherlands: Berkhout Testmateriaal.

Caglar-Ryeng, Ø., Eklund, K., & Nergård-Nilssen, T. (2019). Lexical and grammatical development in children at family risk of dyslexia from early childhood to school entry: A cross-lagged analysis. *Journal of Child Language, 46*, 1102–1126.

Cristia, A., Seidl, A., Junge, C., Soderstrom, M., & Hagoort, P. (2014). Predicting individual variation in language from infant speech perception measures. *Child Development, 85*, 1330–1345.

Field, A. (2013). *Discovering statistics using IBM SPSS Statistics* (4th ed.). London: Sage.

Gelman, A. (2006). Multilevel (hierarchical) modeling: What it can and cannot do. *Technometrics, 48*, 432–435.

Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness, 5*, 189–211.

Gelman, A., & Tuerlinckx, F. (2000). Type S error rates for classical and Bayesian single and multiple comparison procedures. *Computational Statistics, 15*, 373–390.

Gerrits, E., & de Bree, E. (2009). Early language development of children at familial risk of dyslexia: Speech perception and production. *Journal of Communication Disorders, 42*, 180–194.

Goswami, U. (2000). Phonological representations, reading development and dyslexia: Towards a cross-linguistic theoretical framework. *Dyslexia, 6*, 133–151.

Guttorm, T. K., Leppänen, P. H., Poikkeus, A. M., Eklund, K. M., Lyytinen, P., & Lyytinen, H. (2005). Brain event-related potentials (ERPs) measured at birth predict later language development in children with and without familial risk for dyslexia. *Cortex, 41*, 291–303.

Guttorm, T. K., Leppänen, P. H., Hämäläinen, J. A., Eklund, K. M., & Lyytinen, H. J. (2010). Newborn event-related potentials predict poorer pre-reading skills in children at risk for dyslexia. *Journal of Learning Disabilities, 43*, 391–401.

Guttorm, T. K., Leppänen, P. H., Richardson, U., & Lyytinen, H. (2001). Event-related potentials and consonant differentiation in newborns with familial risk for dyslexia. *Journal of Learning Disabilities, 34*, 534–544.

Hakvoort, B., de Bree, E., van der Leij, A., Maassen, B., van Setten, E., Maurits, N., & van Zuijen, T. L. (2016). The role of categorical speech perception and phonological processing in familial risk children with and without dyslexia. *Journal of Speech, Language, and Hearing Research, 59*, 1448–1460.

Hazan, V., Messaoud-Galusi, S., Rosen, S., Nouwens, S., & Shakespeare, B. (2009). Speech perception abilities of adults with dyslexia: Is there any evidence for a true deficit? *Journal of Speech, Language, and Hearing Research, 52*, 1510–1529.

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

Kachel, G., Hardecker, D. J. K., & Bohn, M. (2021). Young children's developing ability to integrate gestural and emotional cues. *Journal of Experimental Child Psychology, 201*, 104984.

Kerkhoff, A., de Bree, E., de Klerk, M., & Wijnen, F. (2013). Non-adjacent dependency learning in infants at familial risk of dyslexia. *Journal of Child Language, 40*, 11–28.

Klerk, M., de, Bree, E., de, Kerkhoff, A., & Wijnen, F. (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. https://doi.org/10.1080/15475441.2018.1497490.

Klerk, M., de, Veen, D., Wijnen, F., & Bree, E. de, (2019). A step forward: Bayesian hierarchical modelling as a tool in assessment of individual discrimination performance. *Infant Behaviour and Development, 57*, 101345. https://doi.org/10.1016/j.infbeh.2019.101345.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*, 831–843.

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 B: Biological Sciences, 363*, 979–1000.

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

Kuijpers, C. T. L., Leij, A., Been, P. H., van Leeuwen, T. G., Keurs, M. T., & Schreuder, R. (2003). Leesproblemen in het voortgezet onderwijs en de volwassenheid. *Pedagogische Studiën, 80*, 272–287.

Leppänen, P. H., Pihko, E., Eklund, K. M., & Lyytinen, H. (1999). Cortical responses of infants with and without a genetic risk for dyslexia: II. Group effects. *NeuroReport, 10*, 969–973.

Leppänen, P. H., Richardson, U., Pihko, E., Eklund, K. M., Guttorm, T. K., Aro, M., & Lyytinen, H. (2002). Brain responses to changes in speech sound durations differ between infants with and without familial risk for dyslexia. *Developmental Neuropsychology, 22*, 407–422.

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, 20*, 335–345.

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 the Acoustical Society of America, 94*, 1242–1255.

Livingston, E. M., Siegel, L. S., & Ribary, U. (2018). Developmental dyslexia: Emotional impact and consequences. *Australian Journal of Learning Difficulties, 23*, 107–135.

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia, 53*, 1–14.

Maassen, B., Groenen, P., Crul, T., Assman-Hulsmans, C., & Gabreëls, F. (2001). Identification and discrimination of voicing and place-of- articulation in developmental dyslexia. *Clinical Linguistics & Phonetics, 15*, 319–339.

Maurer, D., & Werker, J. F. (2014). Perceptual narrowing during infancy: A comparison of language and faces. *Developmental Psychobiology, 56*, 154–178.

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

Messaoud-Galusi, S., Hazan, V., & Rosen, S. (2011). Investigating speech perception in children with dyslexia: Is there evidence of a consistent deficit in individuals?. *Journal of Speech, Language, and Hearing Research, 54*, 1682–1701.

Mittag, M., Thesleff, P., Laasonen, M., & Kujala, T. (2013). The neurophysiological basis of the integration of written and heard syllables in dyslexic adults. *Clinical Neurophysiology, 124*, 315–326.

Mody, M., Student-Kennedy, M., & Brady, S. (1997). Speech Perception Deficits in Poor Readers: Auditory Processing or Phonological Coding? *Journal of Experimental Child Psychology, 64*(2), 199–231. https://doi.org/10.1006/jecp.1996.2343.

Molfese, D. L. (2000). Predicting dyslexia at 8 years of age using neonatal brain responses. *Brain and Language, 72*, 238–245.

Moll, K., Loff, A., & Snowling, M. J. (2013). Cognitive endophenotypes of dyslexia. *Scientific Studies of Reading, 17*, 385–397.

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*, 407–420.

Newman, R., Ratner, N. B., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: A retrospective analysis. *Developmental Psychology, 42*, 643–655.

Nittrouer, S., Shune, S., & Lowenstein, J. H. (2011). What is the deficit in phonological processing deficits: Auditory sensitivity, masking, or category formation?. *Journal of Experimental Child Psychology, 108*, 762–785.

Noordenbos, M. W., Segers, E., Serniclaes, W., Mitterer, H., & Verhoeven, L. (2012). Allophonic mode of speech perception in Dutch children at risk for dyslexia: A longitudinal study. *Research in Developmental Disabilities, 33*, 1469–1483.

Noordenbos, M. W., & Serniclaes, W. (2015). The categorical perception deficit in dyslexia: A meta-analysis. *Scientific Studies of Reading, 19*, 340–359.

Pauls, F., Macha, T., & Petermann, F. (2013). U-shaped development: An old but unsolved problem. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00301.

Pennington, B. F., Santerre-Lemmon, L., Rosenberg, J., MacDonald, B., Boada, R., Friend, A., ... Olson, R. K. (2012). Individual prediction of dyslexia by single versus multiple deficit models. *Journal of Abnormal Psychology, 121*, 212–224.

Peterson, R. L., & Pennington, B. F. (2015). Developmental dyslexia. *Annual Review of Clinical Psychology, 11*, 283–307.

Pihko, E., Leppänen, P. H., Eklund, K. M., Cheour, M., Guttorm, T. K., & Lyytinen, H. (1999). Cortical responses of infants with and without a genetic risk for dyslexia: I. Age effects. *NeuroReport, 10*, 901–905.

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

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., & Frith, U. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain, 126*, 841–865.

Richardson, U., Leppänen, P. H., Leiwo, M., & Lyytinen, H. (2003). Speech perception of infants with high familial risk for dyslexia differ at the age of 6 months. *Developmental Neuropsychology, 23*, 385–397.

Rosen, S., & Manganari, E. (2001). Is there a relationship between speech and nonspeech auditory processing in children with dyslexia?. *Journal of Speech, Language and Hearing Research, 44*, 720–736.

Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science, 12*, 339–349.

Schulte-Körne, G., Deimel, W., Bartling, J., & Remschmidt, H. (2001). Speech perception deficit in dyslexic adults as measured by mismatch negativity (MMN). *International Journal of Psychophysiology, 40*, 77–87.

Serniclaes, W., Van Heghe, S., Mousty, P., Carré, R., & Sprenger-Charolles, L. (2004). Allophonic mode of speech perception in dyslexia. *Journal of Experimental Child Psychology, 87*, 336–361.

Siegler, R. S. (2004). U-shaped interest in U-shaped development—and what it means. *Journal of Cognition and Development, 5*, 1–10.

Snowling, M. J., & Melby-Lervåg, M. (2016). Oral language deficits in familial dyslexia: A meta-analysis and review. *Psychological Bulletin, 142*, 498–545.

Thiede, A., Virtala, P., Ala-Kurikka, I., Partanen, E., Huotilainen, M., Mikkola, K., ... Kujala, T. (2019). An extensive pattern of atypical neural speech-sound discrimination in newborns at risk of dyslexia. *Clinical Neurophysiology, 130*, 634–646.

Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development, 75*, 1067–1084.

Tsuji, S., & Cristia, A. (2014). Perceptual attunement in vowels: A meta-analysis. *Developmental Psychobiology, 56*, 179–191.

Uterwijk, J. (2000). *WAIS-III Nederlandstalige bewerking: Technische handleiding* [WAIS-III Dutch version: Technical manual]. Lisse, Switzerland: Swets & Zeitlinger.

van Alphen, P., de Bree, E., Gerrits, E., de Jong, J., Wilsenach, C., & Wijnen, F. (2004). Early language development in children with a genetic risk of dyslexia. *Dyslexia, 10*, 265–288.

van Bergen, E., de Jong, P. F., Plakas, A., Maassen, B., & van der Leij, A. (2012). Child and parental literacy levels within families with a history of dyslexia. *Journal of Child Psychology and Psychiatry, 53*, 28–36.

van den Bos, K. P., Lutje Spelberg, H., Scheepsma, A., & de Vries, J. (1994). *De Klepel, Vorm A en B: Een test voor de leesvaardigheid van pseudowoorden—Verantwoording, handleiding, diagnostiek en behandeling*. Nijmegen, Netherlands: Berkhout.

van Leeuwen, T., Been, P., Kuijpers, C., Zwarts, F., Maassen, B., & van der Leij, A. (2006). Mismatch response is absent in 2-month-old infants at risk for dyslexia. *NeuroReport, 17*, 351–355.

van Viersen, S., de Bree, E. H., Zee, M., Maassen, B., van der Leij, A., & de Jong, P. F. (2018). Pathways into literacy: The role of early oral language abilities and family risk for dyslexia. *Psychological Science, 29*, 418–428.

van Zuijen, T. L., Plakas, A., Maassen, B. A., Maurits, N. M., & van der Leij, A. (2013). Infant ERPs separate children at risk of dyslexia who become good readers from those who become poor readers. *Developmental Science, 16*, 554–563.

Veenker, T. J. G. (2008). The Zep Experiment Control Application (Version 1.4.1) [computer software]. Utrecht, Netherlands: Utrecht Institute of Linguistics OTS, Utrecht University. Available from http://www.hum.uu.nl/uilots/lab/zep/.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades?. *Journal of Child Psychology and Psychiatry, 45*, 2–40.

Volkmer, S., & Schulte-Körne, G. (2018). Cortical responses to tone and phoneme mismatch as a predictor of dyslexia? A systematic review. *Schizophrenia Research, 191*, 148–160.

Werker, J. F., Hall, D. G., & Fais, L. (2004). Reconstruing U-shaped functions. *Journal of Cognition and Development, 5*, 147–151.

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

Werker, J. F., & Tees, R. C. (1987). Speech perception in severely disabled and average reading children. *Canadian Journal of Psychology/Revue Canadienne de Psychologie, 41*, 48–61.

Wijnen, F. N. K. (2013). Acquisition of linguistic categories: Cross-domain convergences. In J. Bolhuis & M. Everaert (Eds.), *Birdsong, speech and language: Converging mechanisms* (pp. 157–178). Cambridge, MA: MIT Press.