Disentangling attention for frequency and phonological markedness in 9- and 12-month-old infants

Haar, S.M. ter; Levelt, C.C.

Citation

Haar, S. M. ter, & Levelt, C. C. (2018). Disentangling attention for frequency and phonological markedness in 9- and 12-month-old infants. *Language Learning And Development*, *14*(4), 279-296. doi:10.1080/15475441.2018.1480375

Routledge
Taylor & Francis Group

Language Learning and Development

ISSN: 1547-5441 (Print) 1547-3341 (Online) Journal homepage: www.tandfonline.com/journals/hlld20

Disentangling Attention for Frequency and Phonological Markedness in 9- and 12-Month-Old Infants

Sita Minke ter Haar & Clara Cecilia Levelt

To cite this article: Sita Minke ter Haar & Clara Cecilia Levelt (2018) Disentangling Attention for Frequency and Phonological Markedness in 9- and 12-Month-Old Infants, Language Learning and Development, 14:4, 279-296, DOI: 10.1080/15475441.2018.1480375

To link to this article: https://doi.org/10.1080/15475441.2018.1480375

LANGUAGE LEARNING AND DEVELOPMENT
2018, VOL. 14, NO. 4, 279–296
https://doi.org/10.1080/15475441.2018.1480375

Routledge
Taylor & Francis Group

Disentangling Attention for Frequency and Phonological Markedness in 9- and 12-Month-Old Infants

Sita Minke ter Haar [a,b,c] and Clara Cecilia Levelt [a,b]

[a]Leiden Institute for Brain and Cognition, Leiden University, Leiden, the Netherlands; [b]Leiden University Center for Linguist, Leiden University, Leiden, the Netherlands; [c]Institute of Biology Leiden, Leiden University, Leiden, the Netherlands

ABSTRACT
Infants are thought to be sensitive to frequency in the input as a cue for phonological development. However, linguistic biases such as phonological markedness have been argued to play a role too. Since frequency and markedness are correlated, the two assertions could be different interpretations of data that confound frequency and markedness. In this study we disentangle the two, by testing infants' preferences for syllables contrasting in frequency, balanced for markedness, and for syllables contrasting in markedness, balanced for frequency. We expected a developmental change from frequency-independent to frequency-dependent sensitivity. In addition, we expected an early preference for unmarked structure and a later preference for marked structures, as previously found in child language productions. We find that 9-month-olds prefer phonologically unmarked homorganic patterns, independent of frequency, whereas 12-month-olds prefer marked heterorganic patterns. In contrast to what we expected, only a weak effect for frequency is found and no developmental change.

Introduction

Infants have been shown time and again to be sensitive to statistical cues in their language input, like the frequency distribution of sounds and transitional probabilities of syllables (Archer & Curtin, 2011; Kuhl, 1991; Maye, Werker, & Gerken, 2002; Saffran, Aslin, & Newport, 1996; Zamuner, 2006, 2009) but they have also been shown to prefer certain linguistic structures over others, like syllables with a raising sonority pattern over syllables with a falling sonority pattern (Gómez et al., 2014), and place-assimilated consonants in adjacent syllables (oN-Pa: oMPa) over non-place-assimilated consonants (oN-Pa: oNPa) (Smolensky, Jusczyk, & Allocco, 2002). In addition, asymmetries in infants' sensitivity to feature changes have been shown. For example, infants turned out to be less sensitive to Coronal-to-Labial changes than to Labial-to-Coronal changes (Van Der Feest, 2007), and less sensitive to stop-to-continuant changes than to continuant-to-stop changes (Altvater-Mackensen, 2010). Interestingly, most findings showing infants' sensitivity to experience-dependent frequency and statistical cues come from perception experiments performed by psychologists, while most findings showing a role for experience-independent linguistic biases come from (longitudinal) analyses of production data, performed by linguists. Ideally, like in work by Johnson and Jusczyk (2001), we should thus be aware of both types of sensitivity and focus on showing where and when a non-linguistic, statistical factor like frequency is a dominant cue and where and when linguistic biases play a role. An example of such a study is one by Nam and Polka (2016), showing a perceptual bias for stops over fricatives independent of language experience.

CONTACT Sita Minke ter Haar S.M.terHaar@uu.nl Universiteit Leiden, Leiden Institute for Brain and Cognition, Leiden, The Netherlands.

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

The combination of experience-dependent and experience-independent sensitivities becomes all the more important when we try to understand the link between perception and production in acquisition. Do preferences in perception predict the linguistic structure of early productions? Or do young children prefer to listen to linguistic structures that they are able to produce themselves? In a first study that tackled the link between perception and production directly, DePaolis, Vihman, and Keren-Portnoy (2011) found that infants' processing of speech is indeed influenced by their own production capacities.

Here we do not address the link between perception and production directly, but we aim to increase our understanding of early perceptual preferences. By presenting infants with stimuli that can differentiate between the impact of frequency and linguistic biases that we find in early production, we test if perceptual preferences precede production preferences. Jusczyk, Luce, and Charles-Luce (1994) showed a preference for highly frequent CVC syllables over low-frequent CVC syllables in a perception study with 9-month-old infants. However, it is unclear whether it is indeed purely the difference in frequency that determined the preference, or whether linguistic factors played a role as well. Frequency and linguistic well-formedness are not completely independent of each other, as linguistically "unmarked" structures (see below) often occur more frequently in a language than linguistically "marked" structures. It can therefore be hard to separate the two influences from each other. The high- and low-frequency CVC stimuli that were used in Jusczyk et al. (1994) were balanced only for vowel quality, not for other phonological characteristics, and this was done only in one of their experiments. In the experiment where vowel quality was balanced between conditions, the preference for highly frequent patterns was weaker than in the non-balanced experiments, indicating a possible role for phonological biases. In order to test this hypothesis, it would be necessary to pit frequency and linguistic well-formedness against each other in a systematic and controlled way in the stimuli presented to the infants.

This is what was aimed for in the current study, where we tested the perceptual preferences of 9- and 12-month-old infants for lists of carefully constructed consonant-vowel (CV) syllables. We wanted to know whether they would show a preference for more frequent over less frequent syllables in their native language, when balanced for linguistic well-formedness, or for syllables with a specific linguistic well-formedness, when balanced for frequency. If a perceptual preference predicts the structure of early productions (unmarked before marked), we expect to find a preference for linguistically well-formed, i.e., unmarked, syllables, especially in the 9-month-olds. The 12-month-olds might show a development toward less well-formed, i.e., more marked (see below) structures in their perceptual preferences, anticipating the appearance of more marked structures in production. However, if early production abilities affect the perceptual preference we expect the preference for well-formed syllables especially in the 12-month-olds. If we find that there is indeed a clear preference for frequently occurring syllables, independent of linguistic biases, we will need to go back to the early production data in order to assess the role of frequency there. Alternatively, 9-month-olds could pay more attention to general linguistic information (well-formedness) and 12-month-olds more to language-specific information (frequency). If frequency is the main predictor for a preference, we expect a perceptual preference for high-frequency syllables in both the 9- and the 12-month-olds, given the findings by Jusczyk et al. (1994) for 9-month-olds, and the fact that 12-month-olds have gained more distributional experience.

In addition, we were curious if there were any gender differences in linguistic or statistical processing at this age. Clear predictions about the direction of effect are difficult, as the literature is inconclusive in this respect. Related to frequency, Shi (2007) found that while both 8-month-old Quebec-French boys and girls could segment function words, which are highly frequent, from speech, boys showed this with a novelty preference while girls had a familiarity preference. Looking time differences between boys and girls may thus indicate different processing mechanisms rather than some advantage for boys or for girls. However, gender differences are not always found. In the Wordbank database we find a slight developmental advantage for girls over boys in vocabulary development at 9 months, and a larger advantage at 12 months for American English but not for

Danish (Bleses et al., 2008; Eriksson et al., 2012; Fenson et al., 2007; Frank, Braginsky, Yurovsky, & Marchman, 2016)—there was no information in Wordbank for Dutch. Moreover, Huttenlocher et al. suggest potential age-dependent effects of gender and frequency on vocabulary growth (Huttenlocher, Haight, Bryk, Seltzer, & et al, 1991). In our study, boys and girls could differ in their direction of preference, or in the age at which preferred patterns emerge.

In the following sections we describe the current state of the literature on frequency effects and linguistic biases in more detail.

Linguistic biases

The linguistic characteristics that we contrasted in our stimuli were based on what is known about phonological development in child language productions. Regarding the acquisition of classes of sounds, stops are acquired before continuants, voiceless (unaspirated) obstruents before voiced obstruents, Coronal sounds before Labial and Dorsal sounds, and high and low vowels before mid vowels—see examples below (Ingram, 1974b; Jakobson, 1941; McLeod, 2007). These early-acquired classes are usually also more commonly present cross-linguistically (Hyman, 2007; Maddieson, 1991). In phonological theory they are referred to as being “unmarked” or phonologically well-formed, contrasting with “marked” classes that are more unusual cross-linguistically and are acquired later. In addition, in early word productions of Dutch children it had been observed that consonants and vowels within words initially share their place-of-articulation, i.e., early words either consist of combinations of labial consonants and round vowels, or of combinations of coronal consonants and front vowels.[1] The low vowel/a/is neutral and can appear in combination with either coronal consonants or labial consonants. These early forms are termed “homorganic” (Fikkert & Levelt, 2008; Levelt, 1994). Similar homorganic consonant-vowel combinations have been shown cross-linguistically in infant babbling and to a certain extent in adult speech (Davis, MacNeilage, Matyear, & Powell, 2000; MacNeilage, Davis, Kinney, & Matyear, 2000). For ease of exposition we will refer to both the early-acquired classes of sounds and the early-acquired homorganic patterns as “unmarked” and contrast these to “marked” later-acquired sound classes and heterorganic patterns. The phonological characteristics contrasted in this study are described below.

Place of articulation (PoA)

For PoA, Coronal is considered to be the unmarked feature. Cross-linguistic data, language change and production errors all point towards a “special status” for Coronals (Paradis & Prunet, 1991). In early child language productions, Labial and Dorsal consonants are often replaced by Coronal consonants in non-assimilatory contexts, while in the child-specific Consonant Harmony process Coronal consonants are always the target of assimilation (Paradis & Prunet, 1991; in that volume specifically: Lahiri & Evers, 1991; Stemberger & Stoel-Gammon, 1991; Levelt, 1994; (Dutch); Lee, Davis, & Macneilage, 2010). An example of Dorsal consonants being replaced by Coronal consonants in child language is below in (1):

(1) Dorsal → Coronal (Ruth Hills 2;0, (Ingram, 1974a))

Target	Child Production
kiss	[tɪ]
key	[ti:]
go	[dou]

---

[1] A third possible combination is Dorsal consonant + back vowel, but since Dorsal consonants are rare in early productions these patterns are less common, and will therefore not be taken into account here.

Voicing

When studying transcriptions of early English and Dutch child language productions it can be confusing to note that the initial productions of target voiceless plosives by English-speaking babies are transcribed as [b], [d], [g], while those produced by Dutch-speaking babies are transcribed [p], [t], and [k]. Moreover, the initial productions of target voiced plosives (/b/and/d/) by Dutch-speaking babies are transcribed [p] and [t]. This stems from the fact that English is an aspiration language, while Dutch is a pre-voicing language. In both English and Dutch child language, stops are usually first produced with a short-lag, positive VOT value, denoting phonologically [-voice] consonants in Dutch, transcribed/p/,/t/,/k/, but denoting phonologically [-spread glottis] consonants in English, transcribed as/b/,/d/,/g/(Kager, Van Der Feest, Fikkert, Kerkhoff, & Zamuner, 2007). Since our participants are Dutch, the contrast we are interested in is, thus, in terms of [±voice], whereby cross-linguistically [-voice] (expressed by the short-lag, positive VOT) is considered to be the unmarked feature (Van Der Feest, 2007).

(2) Voiced → Voiceless (1;10– 2;1,)

Target	Adult Production	Child Production
douche (‘shower’)	/duʃ/	[tus]
bus (‘bus’)	/bʏs/	[pʏs]
bootje (‘little boat’)	/botjə/	[potjə]

Vowel height

Both [+low], i.e./a/, and [+high], i.e./i/, vowels are considered to be unmarked. Mid vowels, [-high, -low], like/e/,/o/, imply the presence of their low and high counterparts, and are considered marked (Jakobson, 1941). Both low and high vowels appear early in child word productions and with a relatively low error rate, while mid vowels appear later and have a relatively high error rate (Levelt, 1994).

(3) Mid vowel → low/high vowel (Levelt, 1994)

Target	Adult Production	Child Production
beer (‘bear’)	/beɹ/	[biː] (Tirza 1;8.5)
Ernie (name)	/ɛɹni/	[nanɑ] (Jarmo 1;8.12)
pop (‘doll’)	/pɔp/	[pup] (Noortje 2;2.21)

Homorganic or co-occurrence patterns

A pattern on the syllabic level that has been observed in early child language productions is the tendency to produce homorganic syllables, where consonants and vowels in the utterance share their place of articulation (PoA), also referred to as CV co-occurrence patterns (Davis & MacNeilage, 1995; Fikkert & Levelt, 2008; Levelt, 1994; MacNeilage et al., 2000) or consonant-vowel harmony (Seidl & Buckley, 2005). For vowels, Place of Articulation is defined by deviations of the lips and tongue from a neutral, resting state. Labial vowels are those with lip rounding, such as/o/and/u/, dorsal (back) vowels are those with a retracted tongue body, such as/ɑ/,/o/and/u/, and in front vowels such as/i/the tongue tip is close to the teeth. Thus, homorganic syllables require a minimal movement of the articulators between consonant and vowel. As mentioned above, here the homorganic patterns are called “unmarked”.

(4) Homorganic or Co-occurrence patterns (Eva 1;6, (Levelt, 1994))

Target	Adult production	Child Production
brood (‘bread’)	/brot/	[bop]
schoen (‘shoe’)	/sxun/	[pum]
bed (‘bed’)	/bɛt/	[dɛt]

We tested infants’ listening preferences for the above contrasts or for frequency separately, in pairwise stimuli, always balancing the stimuli for all other characteristics as much as possible. This way we were able to disentangle the different factors that influenced—or did not influence—the infants’ perceptual preferences.

Method

Design

We tested Dutch infants’ listening preferences for lists of syllables using a visual fixation method. The lists contrasted in either frequency or markedness. We used a mixed design with age and gender as between-subject factors and all frequency and markedness conditions were tested within each subject.

Participants

Dutch 9- and 12-month-old infants from a monolingual background were tested (9-month-olds: N = 40, 21 boys, 19 girls, mean age 9.02 months (8.45 – 9.50); 12 month-olds N = 24, 13 boys, 11 girls, mean age 12.08 months (11.54 – 12.69). Caregivers reported that the infants developed normally and had no neurological or auditory problems. Nineteen additional 9-month-olds and 21 additional 12-month-olds were tested but were excluded from further analyses because they did not complete the test (9m: N = 6, 12m: N = 7), because they were more than 3 weeks preterm (9m: N = 1, 12m: N = 6) or postterm (12m: N =1), because of dyslexia in the family (9m; N = 2, 12m: N = 1), because of experimental problems and parent interference (9m: N = 3, 12m: N = 4). Infants receiving bilingual linguistic input in their home environment were excluded because this could affect the relative frequencies of sounds to which they are exposed (9m: N = 3, 12m: N = 3). Three infants were excluded because of a combination of factors (9m: N = 3). Individual trials were rejected from the results when they were not reliable (see the statistics section for details). The study was approved by the ethical committee of Leiden University and all caregivers gave written consent for the infants to participate in this study. After the experiment, the caregivers could choose a toy or a book for their participating infant as a token of appreciation.

Stimuli

For the 9-month-olds, two sets of natural stimuli were used; one spoken by a female voice and one by a male voice (21 and 19 infants tested per voice, respectively). The reason we chose two sets of stimuli was to verify that the possible effects were independent of basic acoustic features like voice characteristics and that the results were replicable. Because the statistical model revealed that stimulus voice did not explain any of the variance in the data (indicating it is not an important factor in the model to explain the data).[2]

---

[2] Unfortunately, the lab moved after the experiment on 12-month-olds with stimuli constructed from the female voice, and the 12-months-old infants could therefore not be tested in the male voice condition under the same conditions.

All stimuli were pronounced in a monotonous way and were recorded in a sound attenuated room using Adobe Audition (version 1.5, build 4124.1) and a Sennheiser mkh 416t microphone. Ten sets of Dutch CV syllables were constructed, pairs of which contrasted in frequency (high/low) or markedness (marked/unmarked segments, or homorganic/heterorganic syllables). All stimuli are presented in Appendix Table A2 and A3. All consonants were stops, since these appear early in child language productions (De Boysson-Bardies & Vihman, 1991). Each 10-syllable test trial consisted of 5 different syllables, which were pseudorandomly ordered in two blocks of 5, using Praat (Boersma & Weenink, 2009, version 5.1.25). The syllables were separated by 500ms of silence. The sequences of 10 syllables were presented auditorily while a (motionless) checkerboard pattern was shown on a screen. The sets of syllables were presented in two different orders. There were no large differences in syllable duration or fundamental frequency in any of the stimuli of interest (see Appendix Table A1) and stimuli were rms-equalized. A trained phonetician and a trained phonologist listened to the stimuli and judged them to be representative Dutch syllables.

Frequency: Two sets contrasted in frequency. The syllable frequency was based on an infant directed speech corpus by Van de Weijer (1998). Frequencies were calculated for the stressed CV syllables in infant directed speech, because we only presented them with single CV syllables containing full vowels, and unstressed CV syllables in Dutch usually contain reduced vowels (Booij, 1995). Furthermore, it has been shown that 8- and 9-month-old infants learning a stress-based language such as Dutch, English, or German (and in some cases Spanish/Catalan) are sensitive to syllable stress (Johnson & Jusczyk, 2001; Jusczyk, Houston, & Newsome, 1999) and unstressed, reduced syllables are often deleted in child language (Allen & Hawkins, 1978; Dodd, Holm, Hua, & Crosbie, 2003; Fikkert, 1994; James, Van Doorn, & Mcleod, 2007). The mean frequency values of the frequent and infrequent set are .19% and .01%, respectively, for syllable frequency, and .30% and .05%, respectively, for biphone frequency (relative to the total number of syllables in the database; see Appendix Table A2 and A3). Biphone frequency, the frequency measure used by Jusczyk et al. (1994), is the frequency with which two segments co-occur in all syllables (thus including CVC syllables for example, in contrast to syllable frequency for which we count only CV syllables[3]).

Phonological markedness: The sets of stimuli contrasting in phonological markedness were constructed on the basis of phonological features as described in the Introduction.

Voicing: For the [±voice] condition, the consonants in the two sets of syllables contrasted in being either voiced [+voice], (/b/and/d/), or voiceless (unaspirated) [-voice], (/p/and/t/). The voiceless consonants formed the unmarked category.

Place of Articulation (PoA): Syllable sets in the PoA condition contrasted marked labials, like/p/, with unmarked coronals, like/t/. We did not use dorsal consonants in the sets, because they have a relatively low frequency in Dutch, and we wanted to avoid a potential confound with frequency.

Vowel height: Two sets of syllables contrasted in vowel height. Both [+low], (/a/), and [+high], (/i/), vowels are considered to be unmarked, while the mid vowels, [-high, -low], (/e/,/o/), are marked.

Homorganicity: Two sets contrasted at the level of the entire syllable. The consonant and vowel were homorganic in one set and heterorganic in the other. In case they were homorganic, the entire syllable could be labeled [labial], like/po/with a labial consonant and a labial (i.e., round) vowel, or [coronal], like/ti/with a coronal consonant and a coronal (i.e., front) vowel. In heterorganic syllables the consonant and the vowel carried different PoA features, like/pi/(labial consonant, coronal vowel) or/to/(coronal consonant, labial vowel). Homorganic structures were considered to be phonologically unmarked.

---

[3] There is evidence for infants' sensitivity to syllable weight (Pons & Bosch, 2010; Turk, Jusczyk, & Gerken, 1995) and syllable types (Jusczyk, Jusczyk, Kennedy, Schomberg, & Koenig, 1995). In addition, apart from markedness, syllable type and token frequencies in the input appear to influence the development of syllable production in young children (Jarosz, Calamaro, & Zentz, 2016; Levelt, Schiller, & Levelt, 1999).

Balancing

Care was taken to balance all the sets contrasting in markedness for frequency, and the sets contrasting in frequency for markedness. Since segments can be marked or unmarked in several different ways, stimuli contrasting for one aspect were balanced for all other aspects. For instance, sets contrasting in voicing were balanced for frequency, PoA of the consonant, homorganicity, and vowel height. In the end, the sets of syllables that we compared only clearly contrasted in a single aspect, in order to disentangle the effects of each condition (see Appendix Table A2 and A3).

Apparatus

The experiment was performed in a sound attenuated booth. A chair was placed approximately 90 cm from a 104 cm Philips flat-screen. The stimuli were played in stereo from speakers on both sides of the screen. The screen was connected to a computer outside the booth. Habit X software (Cohen, Atkinson, Chaput, 2000) was used to present the stimuli. Under the screen behind a panel, a camera, and a microphone were placed to monitor the infant's behavior and eye movements from outside the test booth. The video recordings were used for off-line coding of looking time.

Procedure

An adjusted version of the visual-fixation-based auditory preference method was used (Cooper & Aslin, 1990) to test the infants' listening preferences. During the experiment the infant was seated on the caregiver's lap, in front of the screen presenting the stimuli. Caregivers listened to a mix of classical music and backward speech through headphones, to mask the stimuli. A red blinking light was presented on the screen to catch the infant's attention before each trial started. Each set of syllables (trial) lasted for 10 sec, and was presented auditorily while the infant watched a checkerboard pattern on the screen. Each stimulus condition was presented once, thus there were 10 (different) syllable trials in total. Between each syllable trial, the checkerboard was presented again for 10 sec, but a melody was played in order to avoid habituation. The experiment started with two pre-test trials in which all the syllables were presented once in a random order. This was to avoid a primacy effect and to let the infant get used to the setup. After the pre-test and at the end of the whole experiment, a movie of a fish was presented to monitor general attention (measured as looking time). Trials were presented in a different order for each infant.

Scoring

Total looking time was scored for each trial. A difference in looking time between auditory trials is thought to reflect a difference in attentional preference (Colombo & Bundy, 1981; Kemler Nelson et al., 1995). Video files recorded during the experiment were analyzed frame by frame using ELAN software (Sloetjes & Wittenburg, 2008, version 3.7.2). All analyses were performed off-line by one of three trained scorers who were blind to the stimuli. During each 10-sec trial, the looking behavior of the infant was scored. Looks were scored when the infant looked at the center of the screen, the perceptual direction of the sound source. Total looking time (TL) within the 10-sec trial and duration of the longest look (LL) were used as variables. For a reliability estimate, a subset of all data was scored by one of the other experimenters trained in rescoring. The average Pearson's correlation was 0.9 and the average reliability score (intraclass correlation coefficient) was 0.8 ($p = 0.001$ and $p = 0.0001$, respectively).

Statistics

Analyses with linear mixed models were performed using R (R Core Team, 2015) using package lme4 (Bates, Mächler, Bolker, & Walker, 2015) for each of the five conditions (frequency, place of articulation, voicing, vowel height, and homorganicity). Separate models were run with longest looking time (LL) and total looking time (TL) as response variables. The contrast (homorganicity, PoA, voicing, vowel height, or frequency), gender (male/female) and age of the subject were fixed variables, and subject, stimulus voice and order were included as random factors. Data were log-transformed to avoid heteroscedasticity. This step is necessary to meet model assumptions and thus for the model to be applied properly. P-values were determined by model comparison using the Kenward–Roger approximation (Kenward & Roger, 1997) using the *pbkrtest* package in R (Halekoh & Højsgaard, 2014) as comprehensively described in Judd, Westfall, & Kenny (2012). The rationale behind this approach (which is standard for linear mixed models in R) is to see if a factor is relevant for explaining the data fit to the model. Model comparisons were performed starting with the most complex model down to the most simple model. For interaction effects, Tukey post hoc tests were performed using the lsmeans function in R (Lenth, 2016).

Since a separate model was built for each of the five contrasts and for LL and TL data separately, Hochberg's step-up version of the Bonferroni corrections were applied for multiple testing (Hochberg, 1988); see Appendix Table A4 for an overview of the final models.

Individual trials were excluded when the infant was (temporarily) crying or fussy, when the caregiver interrupted or distracted the infant, when the infant did not look within the first 0.5 sec of the trial or when the infant's eyes were not visible. Data from an infant were not included when more than half of the trials had to be excluded or when general attention had dropped to less than 50% by the end of the experiment compared to the beginning (see procedure). Final number of infants per condition are presented in Table A5 and A6, as well as interquartile range and medians.

Data sharing

Anonymized looking time data and stimuli (audio files) will be shared on Databrary (http://databrary.org, Databrary, 2012), an online digital data library that promotes the sharing and reuse of video data.

Results

Homorganicity

By far the strongest effects were found for homorganicity for which the best fitting model included an interaction between stimulus (homorganic/heterorganic) and age (LL: KR-F = 24.4, $p < 0.001$, TL: KR-F = 18.6, $p < 0.001$; Figure 1). Post hoc tests revealed that looking times were significantly longer for homorganic than for heterorganic stimuli in 9-month-olds (TL: b (Estimate) = −0.40, t-ratio = −4.17, $p < 0.001$, LL: b = −0.64, t-ratio = −5.43, $p < 0.001$), whereas 12-month-olds showed the opposite effect with longer looking times for heterorganic stimuli. (TL: b = 0.31, t-ratio = 2.36, p = 0.021; LL, b = 0.33, t-ratio = 2.08, p = 0.04).

Frequency

In the frequency condition the best fitting model includes a significant two-way interaction between stimulus (frequent/infrequent) and gender (male/female) in LL data (LL: F-KR = 7.4, p = 0.007; Figure 2). In post hoc tests, a significant effect for frequency was found in boys, with longer looking times for frequent stimuli than for infrequent stimuli (LL: b = 0.38, t-ratio = 3.63, $p < 0.001$) but not in girls (LL: b = −0.08, t-ratio: - 0.63, p = 0.53). No interactions or main effects were found in TL data ($p > 0.05$).

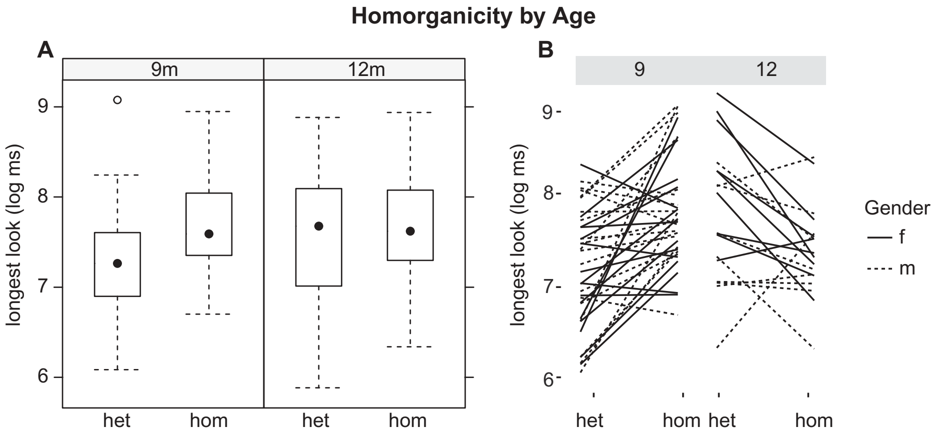

Figure 1. Log of the longest looking time for heterorganic and homorganic stimuli at 9 months (left) and 12 months (right) of age. (**A**) Boxplots indicating group results with interquartile range group medians. (**B**) Individual results. Lines indicate individual data points (solid: female, dotted: male subjects).

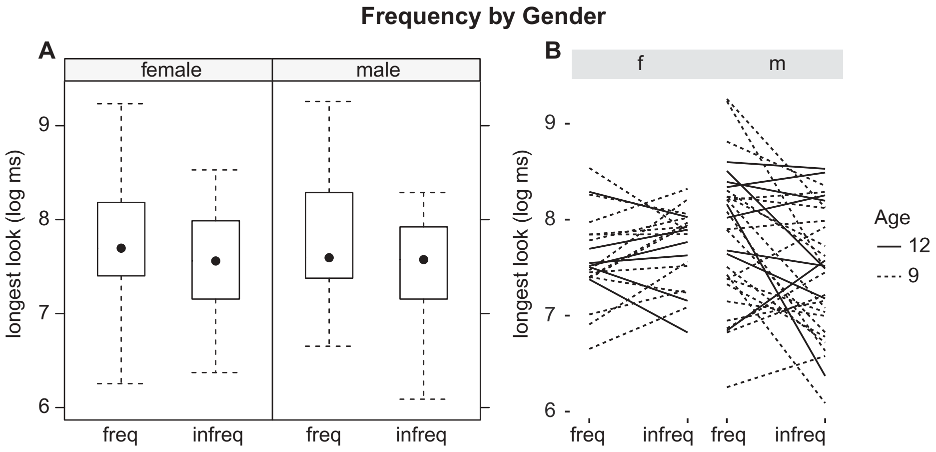

Figure 2. Log of the longest looking time for frequent and infrequent stimuli in girls (left) and boys (right). (**A**) Boxplots indicating group results with interquartile range group medians. (**B**) Individual results. Lines represent individual data points (solid: 12 months, dotted: 9 months).

Vowel height

In LL data, the best fitting model included a two-way interaction between stimulus and gender (LL: F-KR = 5.9, p = 0.02, this it not significant with correction for multiple testing).

Place of articulation

The model best fitting PoA data (both LL and TL) did not include any interactions. Thus, the final model included only the main factor stimulus (coronal/labial). This main factor was not significant in TL data (F-KR = 1.01, p = 0.3). A weak trend for a main effect was found in LL data with slightly longer looking times for labials than for coronals (F-KR = 5.2, p = 0.02, this it not significant with correction for multiple testing).

Voicing

The best fitting model for TL data in the voicing condition included a significant interaction between age and gender ($p = 0.017$). However, this interaction is not meaningful with regard to our study, as the factor stimulus (voiced/voiceless) did not contribute significantly to the model ($p = 0.24$). This means there is no difference in looking time between voiced and voiceless stimuli (only a trend for longer looking times in general for girls than for boys, but this was absent in other conditions).

Discussion

In summary, a clear, age-dependent effect of homorganicity was found, with an early preference for unmarked, homorganic syllables at 9 months of age, which changed toward a preference for marked, heterorganic syllables at 12 months. These effects were visible in both LL and TL data. Interestingly, this developmental path of perceptual preferences fits very well with developmental production data, where a similar change from early homorganic to later heterorganic forms can be found (Levelt, 1994).

Effects of markedness for specific phonological features at the segmental level (PoA, vowel height, and voicing) were less clear. Only a weak indication for a preference for unmarked features was shown, in boys for vowel height. For Place of articulation however, attention for marked stimuli was slightly higher than for unmarked ones. As both these trends were not significant, we cannot draw any conclusions from these data.

In the frequency condition, an attentional preference for frequently occurring syllables was found for boys but not for girls. This effect was found in the LL data only. While it is therefore hard to draw firm conclusions from the frequency data, a possible scenario involving an additive influence of experience-dependent (frequency) and -independent (markedness) effects will be discussed below. We will also reflect on the earlier findings of Jusczyk et al. (1994) with a new perspective on how to interpret their data.

Interestingly, subtle differences between boys and girls appeared in the frequency condition, which we will also shortly discuss below.

Perception and production of homorganic syllables

The findings in the present study point to an early perceptual preference for homorganic syllables, whereas later in development this changes to a preference for heterorganic syllables. At both ages, the infants thus distinguish between the two types of stimuli, however the higher level of attention switches from homorganic to heterorganic syllables. This cannot be explained by a familiarity preference alone, since the stimuli are balanced for frequency.

This pattern is very similar to the development in production, with an early presence of word productions consisting of homorganic syllables, while heterorganic syllables appear only later (see example (4) in the Introduction; Levelt, 1994). The development in production thus appears to follow a similar trajectory as the development in perception we find here. This suggests that linguistic biases at the perceptual level could be predictive of later production skills. However, on the basis of our perception data alone, we cannot exclude the possibility of an influence of production on perception (Davis & MacNeilage, 2000). Homorganic patterns are produced first, and it is possible that during pre-lexical vocal practice, perceptual representations are adjusted to homorganic vocal output. Previous research has shown that an infant's production of certain sound sequences leads to a preference for those sequences in perception (DePaolis et al., 2011; DePaolis, Vihman, & Nakai, 2013; Majorano, Vihman, & DePaolis, 2014). However, in the case of heterorganic syllables, our 12-month-olds show an attentional preference for these syllables before the age they actually produce word forms containing these syllables, pointing to a possible preceding role for perception in this case. This could indicate that for homorganic syllables too, perception foretells production. Alternatively, during the babbling phase, when motor memories are created through production and perceptual feedback (Fagan, 2015), a homorganic bias in both perception and production could develop. Subsequently, by the end of the first year, perceptual attention switches to heterorganic patterns, paving the way for these patterns in production at a somewhat later stage. Longitudinal combined studies of perception and production are needed to provide more insight in how infants' processing of segmental information in perception and production interacts and develops.

Our results also contribute to the idea of an interaction between universal, biological constraints and biases, and experience-dependent learning during development. Homorganic patterns have been shown to be universal across languages, and in babbling and early words, suggesting a possible biological constraint or perceptual bias. The sensitivity for heterorganic patterns at 12 months of age can either be explained by an experience-independent developmental path or by a combination of experience-dependent and -independent mechanisms. That is, due to an initial experience-independent attentional preference, infants first focus on homorganic patterns. Once this stage has passed, a shift to "novel" experience-driven learning occurs, resulting in a preference for heterorganic patterns. Importantly, our homorganic stimuli were balanced for frequency. Therefore, syllable frequency alone cannot explain these results. However, an interaction between frequency and homorganicity is possible. This will be discussed in the section "frequency revisited" below.

Similar developmental patterns have been found in infants' development of musical preferences (Hannon, Soley, & Levine, 2011) as well as in songbirds' vocal learning (Ter Haar, Kaemper, Stam, Levelt, & Ten Cate, 2014). Discrimination of simple beat ratios is unaffected by the music present in the infant's culture early in development, but older infants are affected by the beat patterns in their culture's music. This points to an early experience-independent sensitivity and a later experience-dependent sensitivity (Hannon et al., 2011). Zebra finches (*Taeniopygia guttata*, a vocally learning species) that were raised by their non-singing mother, and therefore had no auditory song experience, showed a perceptual preference for universal zebra finch song elements. After experience with a specific song, the preference of the same birds changed to the (non-universal) song elements they had been exposed to Ter Haar et al. (2014). Thus, interactions between biologically determined constraints and biases, and experience during development seem to exist in all these domains.

Development of segmental representations

No clear attentional preference for either marked or unmarked stimuli was found here for the segmental features (voicing, PoA, and vowel height). No effect at all was found for voicing and for PoA and only trends were found for vowel height. If anything, the trends were in the opposite direction of what we expected. For PoA there was a trend for longer looks for marked stimuli (labial over coronal), whereas a trend for longer looks for unmarked vowels was found, but only in boys.

This lack of strong effects may be explained by the preference for homorganic syllables, which could indicate that young infants perceive feature contrasts at the syllable level and not yet at the segmental level (Ferguson & Farwell, 1975; Fikkert & Levelt, 2008; Levelt, 1994; Waterson, 1971). In production, differentiation of PoA features for consonants and vowels within syllables only appeared after age 1;5 in the Dutch children studied in Fikkert and Levelt (2008). In the present study, the preference for heterorganic syllables is only found in the 12-month-olds. This could indicate that they start to become sensitive to features at the segmental level at this age. A recent study on vowel discrimination in Dutch-learning infants (Liu & Kager, 2016) shows that the vowel contrast (between/I/and/i/) is only discriminated at this same age (11–12 months) in monolinguals, which gives support to our explanation. However, whether this applied to specific contrasts used in the present study should be confirmed. We would also expect that discrimination between homorganic syllables is easier than heterorganic ones if segmental representation is not yet fully developed (i.e., before 12 months of age).

Others have suggested that segmental awareness arises before 12 months of age (Nishibayashi & Nazzi, 2016), since infants have a vowel bias at 6 months of age and a consonant bias at 8 months of age (the latter depending on language input). This means that infants primarily use vowel information at 6 months of age and consonant information at 8 months of age for word identification (measured in a mispronunciation task). However, this was tested in a discrimination paradigm (familiarization followed by mispronunciation) at the syllable level. For consonants, it is nearly impossible to truly determine segmental awareness, since consonantal contrasts are always presented embedded in syllables, and therefore constitute a contrast at the syllable level as well. This means that also in our study, we cannot determine with certainty whether or not infants process the stimuli at the segmental level.

An alternative explanation could lie in the processing of complex information (in the sense of number of features in a syllable), which causes the infants to initially attend to homorganic syllables and later to heterorganic (more complex) syllables. Processing of complex information does not necessarily have to take place at segmental level.

Frequency effects revisited

In 1994, Jusczyk et al. published the first evidence for an effect of frequency in the perceptual preferences of 9-month-old but not 6-month-old infants (Jusczyk et al., 1994). Since then, the frequency effect has been taken for granted, but surprisingly, it has not been reproduced in preference tests other than the current study. However, other indications for frequency effects have been shown (Archer & Curtin, 2011; Kuhl, 1991; Maye et al., 2002; Saffran et al., 1996; Zamuner, 2006, 2009), but very rarely are stimuli controlled for linguistic factors like markedness. The only evidence for a frequency effect on phonological development independent of markedness has been shown by Zamuner, Gerken, and Hammond (2005). In our study we do find a small frequency effect when balancing for markedness and homorganicity, but it is much weaker than the effect for homorganicity and only significant for boys, not for girls. It would be interesting to know if this gender difference was also present in the study by Jusczyk et al. (1994) but gender appears not to have been investigated as a factor in that study.

So why was the frequency effect so strong in Jusczyk et al. (1994)? In the first two experiments in that study, phonological factors were not controlled for at all in the stimuli. In the third experiment, the frequent and infrequent lists of syllables were balanced for vowel quality and the effect was already much weaker. However, they did not balance the consonants for markedness in the third experiment, and the frequent syllables lists had a higher numbers of initial plosives (41 items) than initial fricatives (77 items) compared to infrequent lists of syllables (10 plosives and 126 fricatives). Importantly, infrequent lists of syllables contained 35 items with the voiced dental non-sibilant fricative [ð] whereas frequent lists contained none. Similarly, frequent lists contained 10 alveolar approximants [ɹ] whereas the infrequent lists contained none. This means that just by an increase in preference for [ð] or a lack of preference for [ɹ] could explain the results reported by Jusczyk et al. (1994). Furthermore, even though overall, the lists were equal in homorganicity, only the frequent list contained syllables starting with plosives that were homorganic (10 items), whereas in the infrequent list no homorganic syllables starting with plosives occurred at all. It is possible that infants will pay more attention to syllables with initial plosives than to syllables with initial fricatives and even more so when they are homorganic. Strictly speaking, different distribution of homorganic syllables starting with a plosive in the two sets might be sufficient to explain the difference in preference between “high frequency” and “low frequency” lists in Jusczyk et al., and would be highly similar to the homorganicity preference we show for plosives in our present study. Linguistic factors could thus have influenced the looking times for the frequent syllables, and the combined information might have triggered the significantly longer looking times for frequent syllables. Thus, it seems plausible that the combination of phonological information and frequency, rather than frequency alone explain the perceptual preferences in Jusczyk et al. (1994). It should be noted that the frequency difference in Jusczyk et al. (1994) was larger than the frequency difference in our study, which is a consequence of our balancing for markedness. For positional phoneme frequency, the frequent/infrequent ratio in Jusczyk et al. (experiment 3, which is balanced for vowel quality only) is 2.8, while it is 1.1 in the present study. In Jusczyk et al., the frequent/infrequent ratio for biphone frequency was 5.8, compared to 6 in the present study and the syllable frequency difference ratio here is 16.7 (the syllable frequency difference ratio is not reported in Jusczyk et al). By no means do we exclude a role for frequency in infants' perceptual preferences. However, our results point out that other factors, such as homorganicity can at least additionally, but possibly even more strongly affect perceptual preference.

Gender differences

We found gender differences in the attentional preferences for frequency. The frequency preference was only present in boys, alluding to differences in perceptual biases between boys and girls. However, it is unclear how this gender difference could relate to either the vocabulary advantage for American English acquiring girls, or the novelty—as opposed to a female familiarity-preference for function words found in Shi (2007). Reports on gender differences in language development are sparse and often conflicting. Many different causes can be attributed to gender differences, such as (social) environment, parental speech, or biological factors, which are not mutually exclusive. For example, indications that differences in gender and/or hormone levels are related to language development are accumulating (Schaadt, Hesse, & Friederici, 2015). In future work, the possible this difference in sensitivity to frequency between boys and girls, all else being equal, could be tested in different linguistic domains. Our finding is yet another indication that it is relevant to take potential gender differences into account.

Conclusion

In summary, our results suggest that the homorganic patterns found in early word productions are preceded by a perceptual sensitivity to homorganic syllables. Preferences in perception thus seem to be predictive of the linguistic structure of early productions. Homorganicity turns out to be a strong perceptual attractor early in development, and possibly even a stronger cue than sheer frequency. We also showed that it is important to control for both frequency and linguistic characteristics of stimuli in studies testing perceptual sensitivity, as the two factors are often confounded. Phonological development is influenced by both linguistic and non-linguistic factors, and our knowledge of early language development benefits from interdisciplinary research that their relative role, and explains the ways they interact.

Acknowledgments

We thank Monique Bisschop and Anne-Marie Roorda for help with scoring the videos and we thank all the babies and their caregivers for their participation in this study.

Funding

This work was supported by the Leiden Institute for Brain and Cognition.

ORCID

Sita Minke ter Haar http://orcid.org/0000-0002-7408-8864
Clara Cecilia Levelt http://orcid.org/0000-0003-2916-0326

References

Allen, G. D., & Hawkins, S. (1978). The development of phonological rhythm. In A. Bell & J. B. Hooper (Eds.), *Syllables and segments*. New York: Elsevier-North Holland.

Altvater-Mackensen, N. (2010). *Do manners matter? Asymmetries in the acquisition of manner of articulation features*. Nijmegen: Radboud Universiteit Nijmegen.

Archer, S. L., & Curtin, S. (2011). Perceiving onset clusters in infancy. *Infant Behavior & Development*, *34*(4), 534–540. doi:10.1016/j.infbeh.2011.07.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. Eprint arXiv:1406.5823. *Journal of Statistical Software*, *67*(1), 51. doi:10.18637/jss.v067.i01

Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T. O., & Basbøll, H. (2008). The danish communicative developmental inventories: Validity and main developmental trends. *Journal of Child Language*, *35*(3), 651–669. doi:10.1017/S0305000907008574

Boersma, P., & Weenink, D. (2009). Praat: doing phonetics by computer [Computer Program] (Version 5.1.25).

Booij, G. E. (1995). The phonology of Dutch. Oxford University Press. Retrieved from https://books.google.nl/books/about/The_Phonology_of_Dutch.html?id=LT6E6YdAh-MC&redir_esc=y

Cohen, L. B., Atkinson, D. J., & Chaput, H. (2000). Habit 2000: A new program for testing infant perception and cognition (version 2. 2. 5c) [computer software]. Austin: University of Texas.

Colombo, J., & Bundy, R. S. (1981). A method for the measurement of infant auditory selectivity. *Infant Behavior and Development*, *4*, 219–223. doi:10.1016/S0163-6383(81)80025-2

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, *61*(5), 1584–1595. doi:10.1111/j.1467-8624.1990.tb02885.x

Databrary. (2012). The Databrary Project: A video data library for developmental science. New York: New York University. Retrieved from http://databrary.org.

Davis, B. L., & MacNeilage, P. F. (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research*, *38*(6), 1199–1211. doi:10.1044/jshr.3806.1199

Davis, B. L., & MacNeilage, P. F. (2000). An embodiment perspective on the acquisition of speech perception. *Phonetica*, *57*(2–4), 229–241. doi:10.1159/000028476

Davis, B. L., MacNeilage, P. F., Matyear, C. L., & Powell, J. K. (2000). Prosodic correlates of stress in babbling: An acoustical study. *Child Development*, *71*(5), 1258–1270. doi:10.1016/S0163-6383(96)90312-4

De Boysson-Bardies, B., & Vihman, M. M. (1991). Adaptation to language: Evidence from babbling and first words in four languages. *Language*, *67*(2), 297–319. doi:10.1353/lan.1991.0045

DePaolis, R. A., Vihman, M. M., & Keren-Portnoy, T. (2011). Do production patterns influence the processing of speech in prelinguistic infants? *Infant Behavior & Development*, *34*(4), 590–601. doi:10.1016/j.infbeh.2011.06.005

DePaolis, R. A., Vihman, M. M., & Nakai, S. (2013). The influence of babbling patterns on the processing of speech. *Infant Behavior and Development*, *36*(4), 642–649. doi:10.1016/j.infbeh.2013.06.007

Dodd, B., Holm, A., Hua, Z., & Crosbie, S. (2003). Phonological development: A normative study of British English-speaking children. *Clinical Linguistics & Phonetics*, *17*(8), 617–643. doi:10.1080/0269920031000111348

Eriksson, M., Marschik, P. B., Tulviste, T., Almgren, M., Pérez Pereira, M., Wehberg, S., … Gallego, C. (2012). Differences between girls and boys in emerging language skills: Evidence from 10 language communities. *British Journal of Developmental Psychology*, *30*(2), 326–343. doi:10.1111/j.2044-835X.2011.02042.x

Fagan, M. K. (2015). Why repetition? Repetitive babbling, auditory feedback, and cochlear implantation. *Journal of Experimental Child Psychology*, *137*, 125–136. doi:10.1016/j.jecp.2015.04.005

Fenson, L., Marchman, V. A., Thal, D., Dale, P., Reznick, J. S., & Bates, E. (2007). *MacArthur-bates communicative development inventories: User's guide and technical manual* (2nd ed.). Baltimore, MD: Brookes Publishing Co.

Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language*, *51*(2), 419–439. Retrieved from http://www.jstor.org/stable/412864

Fikkert, P. (1994). *On the acquisition of prosodic structure*. The Hague: Leiden University. Holland Academic Graphics. Leiden, the Netherlands.

Fikkert, P., & Levelt, C. (2008). *How does place fall into place? contrast in phonology theory, perception, acquisition*. Berlin, Boston: De Gruyter. doi:10.1515/9783110208603.3.231

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2016). Wordbank: An open repository for developmental vocabulary data*. *Journal of Child Language*. Retrieved from http://langcog.stanford.edu/papers_new/frank-2016-jcl.pdf

Gómez, D. M., Berent, I., Benavides-Varela, S., Bion, R. A. H., Cattarossi, L., Nespor, M., & Mehler, J. (2014). Language universals at birth. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(16), 5837–5841. doi:10.1073/pnas.1318261111

Halekoh, U., & Højsgaard, S. (2014). A kenward-roger approximation and parametric bootstrap methods for tests in linear mixed models - The R package pbkrtest. *Journal of Statistical Software*, *59*(9), 1–32. doi:10.18637/jss.v059.i09

Hannon, E. E., Soley, G., & Levine, R. S. (2011). Constraints on infants' musical rhythm perception: Effects of interval ratio complexity and enculturation. *Developmental Science*, *14*(4), 865–872. doi:10.1111/j.1467-7687.2011.01036.x

Hochberg, Y. (1988). A sharper bonferroni procedure for multiple tests of significance. *Biometrika*, *75*(4), 800–801. doi:10.1093/biomet/75.4.800

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236–248. doi:10.1037/0012-1649.27.2.236

Hyman, L. M. (2007). Universals in phonology. *Phonology*, *25*(2007), 98–109. doi:10.1515/TLIR.2008.003

Ingram, D. (1974a). Fronting in child phonology. *Journal of Child Language*, *1*(2), 233–241. doi:10.1017/S0305000900000672

Ingram, D. (1974b). Phonological rules in young children. *Journal of Child Language*, *1*(1), 49–64. doi:10.1017/S0305000900000076

Jakobson, R. (1941). *Child language, aphasia and phonological universals.* (Translated by A. Keiler, 1968). Den Haag-Parijs, SouthHolland, The Netherlands: Mouton.

James, D. G. H., Van Doorn, J., & Mcleod, S. (2007). Does the number of syllables in words affect weak syllable deletion in typically developing children aged 3 to 7 years speaking Australian english? In World congress of the International Association of Logopedia and Phoniatrics (IALP), 27. Copenhagen, Denmark. Retrieved from http://www.sld.cu/galerias/pdf/sitios/rehabilitacion-logo/silabas_.pdf

Jarosz, G., Calamaro, S., & Zentz, J. (2016). Input frequency and the acquisition of syllable structure in Polish. *Language Acquisition*, *24*(4), 361–399. doi:10.1080/10489223.2016.1179743

Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, *44*, 548–567. doi:10.1006/jmla.2000.2755

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. doi:10.1037/a0028347

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in english-learning infants. *Cognitive Psychology*, *39*(3–4), 159–207. doi:10.1006/cogp.1999.0716

Jusczyk, P. W., Jusczyk, A. M., Kennedy, L. J., Schomberg, T., & Koenig, N. (1995). Young infants' retention of information about bisyllabic utterances. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(4), 822–836. doi:10.1037/0096-1523.21.4.822

Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994, October). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, *33*(5), 630. doi:10.1006/jmla.1994.1030

Kager, R., Van Der Feest, S., Fikkert, P., Kerkhoff, A., & Zamuner, T. S. (2007). Representations of [voice]: Evidence from acquisition. In J. Van De Weijer & E. Van Der Torre (Eds.), *Voicing in Dutch* (pp. 41–80). Amsterdam, the Netherlands.: John Benjamins. doi:10.1075/cilt.286.03kag

Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. (1995). The head-turn preference procedure for testing auditory perception. *Infant Behavior and Development*, *18*(1), 111–116. doi:10.1016/0163-6383(95)90012-8

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, *53*(3), 983–997. Retrieved from http://www.jstor.org/stable/2533558

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, *50*(2), 93–107. doi:10.3758/BF03212211

Lahiri, A., & Evers, V. (1991). Palatalization and coronality. In C. Paradis & J.-F. Prunet (Eds.), *The special status of coronals: Internal and external evidence* (pp. 79–100). Phonology and Phonetics 2. San Diego, USA.: Academic.

Lee, S.A. S., Davis, B., & Macneilage, P. (2010). Universal production patterns and ambient language influences in babbling: a crosslinguistic study of Korean- and English-learning infants. *Journal of Child Language*, *37*(2), 293–318. https://doi.org/10.1017/S0305000909009532

Lenth, R. V. (2016). Least-squares means: The {R} package {lsmeans}. *Journal of Statistical Software*, *69*(1), 1–33. doi:10.18637/jss.v069.i01

Levelt, C. C. (1994). *On the acquisition of place*. Leiden, the Netherlands.: Leiden University.

Levelt, C. C., Schiller, N. O., & Levelt, W. J. M. (1999). A developmental grammar for syllable structure in the production of child language and languages have different syllable inventories. in terms of CV-structure, other languages have an extensive inventory which includes CV, CVC, V. *Brain and Language*, *299*, 291–299. doi:10.1006/brln.1999.2092

Liu, L., & Kager, R. (2016). Perception of a native vowel contrast by Dutch monolingual and bilingual infants: A bilingual perceptual lead. *International Journal of Bilingualism*, *20*(3), 335–345. doi:10.1177/1367006914566082

MacNeilage, P. F., Davis, B. L., Kinney, A., & Matyear, C. L. (2000). The motor core of speech: A comparison of serial organization patterns in infants and languages. *Child Development*, *71*(1), 153–163. doi:10.1111/1467-8624.00129

Maddieson, I. (1991). Testing the Universality of phonological generalizations with a phonetically specified segment database: Results and limitations. *Phonetica*, *48*(2–4), 193–206. doi:10.1159/000261884

Majorano, M., Vihman, M. M., & DePaolis, R. A. (2014). The relationship between infants? production experience and their processing of speech. *Language Learning and Development*, *10*(2), 179–204. doi:10.1080/15475441.2013.829740

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), 101–111. doi:10.1016/S0010-0277(01)00157-3

McLeod, S. (Ed.). (2007). *The international guide to speech acquisition*. Clifton Park, NY: Thomson Delmar Learning.

Nam, Y., & Polka, L. (2016). The phonetic landscape in infant consonant perception is an uneven terrain. *Cognition*, *155*, 57–66. doi:10.1016/j.cognition.2016.06.005

Nishibayashi, L. L., & Nazzi, T. (2016). Vowels, then consonants: Early bias switch in recognizing segmented word forms. *Cognition*, *155*, 188–203. doi:10.1016/j.cognition.2016.07.003

Paradis, C., & Prunet Jean-Francois. (1991). *Phonetics and phonology - the special status of Coronals: Internal and external evidence*. (C. Paradis & PrunetJean-Francois, Eds.). Cambridge, USA: Academic Press.

Pons, F., & Bosch, L. (2010). Stress pattern preference in spanish-learning infants: The role of syllable weight. *Infancy*, *15*(3), 223–245. doi:10.1111/j.1532-7078.2009.00016.x

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(December), 926–928. doi:10.1126/science.274.5294.1926

Schaadt, G., Hesse, V., & Friederici, A. D. (2015). Sex hormones in early infancy seem to predict aspects of later language development. *Brain and Language*, *141*, 70–76. doi:10.1016/j.bandl.2014.11.015

Seidl, A., & Buckley, E. (2005). On the learning of arbitrary phonological rules. *Language Learning and Development*, *1* (3), 289–316. doi:10.1207/s15473341lld0103&4_4

Shi, R. (2007). Infants ' recognition of function words in continuous speech. In ICPhS XVI. (pp. 1541–1544).Saarbrücken Univ. des Saarlandes 2007, Germany.

Sloetjes, H., & Wittenburg, P. (2008). Annotation by category – ELAN and ISO DCR. 6th International Conference on Language Resources and Evaluation (LREC 2008).Proceedings, LREC 2008 : Proceedings 6th International Conference on Language Resources and Evaluation.

Smolensky, P., Jusczyk, P. W., & Allocco, T. (2002). How english-learning infants respond to markedness and faithfulness constraints. *Language Acquisition*, *10*(1), 31–73. doi:10.1207/S15327817LA1001

Stemberger, J. P., & Stoel-Gammon, C. (1991). The Underspecification of Coronals: Evidence From Language Acquisition and Performance Errors. In *The Special Status of Coronals: Internal and External Evidence* (pp. 181–199). Elsevier. https://doi.org/10.1016/B978-0-12-544966-3.50015-4

Ter Haar, S. M., Kaemper, W., Stam, K., Levelt, C. C., & Ten Cate, C. (2014). The interplay of within-species perceptual predispositions and experience during song ontogeny in zebra finches (Taeniopygia guttata). *Proceedings. Biological Sciences*, *281*(1796), 20141860. doi:10.1098/rspb.2014.1860

Turk, A. E., Jusczyk, P. W., & Gerken, L. (1995). Do english-learning infants use syllable weight to determine stress? *Language and Speech*, *38*(2), 143–158. doi:10.1177/002383099503800202

Van De Weijer, J. C. (1998). *Language input for word discovery*. MPI series in psycholinguistics. Radboud University Nijmege, Nijmegen. https://doi.org/doi:10.17617/2.2057670

Van Der Feest, S. (2007). *Building a phonological lexicon: The acquisition of the Dutch voicing contrast in perception and production*. Radboud Universiteit Nijmegen, Nijmegen.

Waterson, N. (1971). Child phonology: A prosodic view. *Journal of Linguistics*, *7*, 179–211. doi:10.1017/S0022226700002917

Zamuner, T. S. (2006). Sensitivity to word-final phonotactics in 9- to 16-month-old infants. *Infancy*, *10*(1), 77–95. doi:10.1207/s15327078in1001_5

Zamuner, T. S. (2009). Phonotactic probabilities at the onset of language development: Speech production and word position. *Journal of Speech, Language, and Hearing Research*, *52*(1), 49–60. doi:10.1044/1092-4388(2008/07-0138)

Zamuner, T. S., Gerken, L., & Hammond, M. (2005). The acquisition of phonology based on input: A closer look at the relation of cross-linguistic and child language data. *Lingua*, *115*(10), 1403–1426. doi:10.1016/j.lingua.2004.06.005

Appendix

Table A1. Acoustic properties of the stimuli.

		Female voice						Male voice					
		F0	min	max	Dur	min	max	F0	min	max	Dur	min	max
PoA	Coronal	200	197	204	388	261	142	133	129	142	377	223	223
	Labial	205	201	209	399	292	141	137	132	141	365	192	466
Voice	Voiceless	203	198	206	337	292	141	136	132	141	311	192	360
	Voiced	203	209	209	444	387	137	134	129	137	412	324	471
Vh	High/Low	203	197	210	355	310	138	134	130	138	353	257	471
	Mid	200	197	203	404	350	143	137	130	143	371	326	420
Hom	Homorganic	201	197	206	362	261	143	137	130	143	352	223	466
	Heterorganic	203	198	209	389	310	138	136	129	138	358	257	428
Freq	Frequent	202	197	209	426	353	143	136	130	143	389	326	420
	Infrequent	204	200	206	368	200	454	134	129	142	355	223	466

Notes. Acoustic properties of the two sets of stimuli; Male voice and Female voice. PoA: place of articulation, Vh: vowel height, Freq: frequency of occurrence, Dur: mean syllable duration in ms, F0: mean fundamental frequency in Hz, range: minimum and maximum F0 and duration, respectively.<?TFN>

Table A2. Stimulus syllables per condition with frequency values for phonological features.

Condition	Unmarked Syllables				Marked Syllables		
PoA	Coronal	f%	bf%		Labial	f%	bf%
	/do/	0.01	0.12		**/bo/**	0.19	0.27
	/da/	0.09	0.74		**/ba/**	0.01	0.02
	/to/	0.03	0.03		**/po/**	0.01	0.02
	/ti/	0.02	0.07		**/pi/**	0.02	0.09
	/de/	0.16	0.19		**/be/**	0.17	0.55
Total		**0.30**	**1.15**	**Total**		**0.40**	**0.94**
Average		**0.06**	**0.23**	**Average**		**0.08**	**0.19**
Voice	Voiceless				Voiced		
		f%	bf%			f%	bf%
	/pi/	0.02	0.09		**/bi/**	0.01	0.01
	/po/	0.01	0.02		**/bo/**	0.19	0.27
	/to/	0.03	0.03		**/do/**	0.01	0.12
	/pe/	0.02	0.03		**/be/**	0.17	0.55
	/ta/	0.32	0.34		**/da/**	0.09	0.74
Total		**0.40**	**0.50**	**Total**		**0.46**	**1.68**
Average		**0.08**	**0.10**	**Average**		**0.09**	**0.34**
Vowel height	Unmarked: i,a (high/low)				Marked: o,e (mid)		
		f%	bf%			f%	bf%
	/da/	0.01	0.74		**/de/**	0.03	0.19
	/ki/	0.02	0.03		**/ke/**	0.03	0.21
	/ka/	0.12	0.21		**/ko/**	0.05	0.14
	/ta/	0.16	0.34		**/to/**	0.17	0.03
	/bi/	0.01	0.01		**/bo/**	0.01	0.27
Total		**0.32**	1.33	**Total**		**0.28**	**0.84**
Average		**0.06**	**0.27**	**Average**		**0.06**	**0.17**
Homorganicity	Unmarked (homorganic)				Marked (heterorganic)		
		f%	bf%			f%	bf%
	/po/	0.09	0.02		**/to/**	0.16	0.03
	/ti/	0.03	0.07		**/ki/**	0.05	0.03
	/ko/	0.15	0.14		**/ke/**	0.12	0.21
	/de/	0.32	0.19		**/be/**	0.03	0.55
	/ba/	0.01	0.02		**/do/**	0.19	0.12
Total		**0.59**	0.44	**Total**		**0.56**	**0.95**
Average		**0.12**	**0.09**	**Average**		**0.11**	**0.19**

Notes. f%: stressed syllable frequency and bf% biphone frequency, both in percentage of total tokens in the Van de Weijer database of infant directed speech in Dutch (Van De Weijer, 1998). Syllable notations are according to International Phonetic Alphabet (IPA). For example,/ba/as in the Dutch word “baken”.

Table A3. Stimulus syllables per condition with frequency values for the frequency condition.

Condition							
Frequency	Frequent Syllables	f%	bf%		Infrequent Syllables	f%	bf%
	Bo	0.19	0.27		**bi**	0.01	0.01
	Be	0.17	0.55		**ba**	0.01	0.02
	De	0.16	0.19		**do**	0.01	0.12
	Ta	0.32	0.34		**ti**	0.02	0.07
	Ko	0.12	0.14		**pe**	0.02	0.03
Total		**0.96**	**1.48**	**Total**		**0.06**	**0.25**
Average		**0.19**	**0.30**	**Average**		**0.01**	**0.05**

Notes: f%: stressed syllable frequency and bf% biphone frequency, both in percentage of total tokens in the Van de Weijer database of infant directed speech in Dutch (Van De Weijer, 1998). Syllable notations according to International Phonetic Alphabet.

Table A4. Overview of final model for each condition.

		TL		LL	
Condition	Optimal model	p-value	F	p-value	F
Homorganicity	Age * Stimulus	< 0.0001 *	18.63	<0.0001 *	24.42
PoA	Stimulus (main effect)	> 0.05	1.02	0.03 #	5.25
Voicing	Stimulus (main effect)	> 0.05	1.44	> 0.05	2.31
Vowel height	Gender * Stimulus	> 0.05	0.04	0.02 #	5.61
Frequency	Gender * Stimulus	> 0.05	0.16	0.007 *	7.76

Notes. Final models selected based on model comparison using Kenward-Roger approximation. All models include subject as a random factor. TL: total looking time; LL: longest look; p-values: * significant with Hochberg's step-up correction, # trend: not significant with Hochberg's step-up correction, ns p > 0.05); F: Ftest statistics from Kenward-Roger approximation.

Table A5. Interquartile range. Median, and N for each condition for total looking times (ms).

	IQR	Median	N	IQR	median	N
Frequent	2334	3870	35	658	4142	15
Infrequent	2162	4107	35	2048	3076	15
PoA UM(coronal)	3047	3370	32	2956	2219	13
PoA M (labial)	3366	3402	32	2524	2991	13
Homorganic	2558	4841	35	2400	3068	18
Heterorganic	2622	3389	35	3593	4700	18
Voiceless	3507	3630	32	3262	3897	19
Voiced	2420	4135	32	3192	4986	19
Vowel UM (high/low)	2326	3680	31	2918	3680	19
Vowel M (mid)	3255	4235	31	1990	4060	19

Table A6. Interquartile range, median, and N for each condition for longest looking times (ms).

	9 mothns			12 months		
	IQR	median	N	IQR	median	N
Frequent	1822	3870	35	2181	4142	15
Infrequent	1840	4107	35	2054	3076	15
PoA UM(coronal)	1297	3370	32	1550	2219	17
PoA M (labial)	1199	3402	32	2089	2991	17
Homorganic	2277	4841	35	1680	3068	18
Heterorganic	1346	3389	35	1948	4700	18
Voiceless	1960	3630	32	2077	3897	19
Voiced	1790	4135	32	2820	4986	19
Vowel UM (high/low)	2016	3680	31	2000	3680	19
Vowel M (mid)	2090	4235	31	2955	4060	19