

Pronunciation Variation in Infant-Directed Speech: Phonetic Reduction of Two Highly Frequent Words

Mybeth Lahey¹

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands;

Centre for Language Studies, Radboud University Nijmegen, Nijmegen, The Netherlands; and

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

Wundtlaan 1,

6525 XD Nijmegen,

The Netherlands

mybeth.lahey@mpi.nl

Mirjam Ernestus

Centre for Language Studies, Radboud University Nijmegen, Nijmegen, The Netherlands; and

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Wundtlaan 1,

6525 XD Nijmegen,

The Netherlands

m.ernestus@let.ru.nl

¹ Author to whom correspondence should be addressed.

Abstract

In spontaneous conversations between adults, words are often pronounced with fewer segments or syllables than their citation forms. The question arises whether infant-directed speech also contains phonetic reduction. If so, infants would be presented with speech input that enables them to acquire reduced variants from an early age. This study compared speech directed at 11- and 12-month-old infants with adult-directed conversational speech and adult-directed read speech. In an acoustic study, 216 tokens of the Dutch words *allemaal* and *helemaal* from speech corpora were analyzed for duration, number of syllables and vowel quality. In a perception study, adult participants rated these same materials for reduction and provided phonetic transcriptions. The results show that these two words are frequently reduced in infant-directed speech, and that their degree of reduction is comparable to conversational adult-directed speech. These findings suggest that lexical representations for reduced pronunciation variants can be acquired early in linguistic development.

An important component of language acquisition is the learning of words and their meanings. Before infants can attach meanings to words, they must acquire the ability to recognize a given word across different contexts. This is a complicated task, as each token of a word is acoustically different. Pronunciation variation can, for instance, be caused by speaker gender, voice characteristics and accent or dialectal variation. This raises the question of how infants learn to cope with pronunciation variation.

In spontaneous conversations between adults, another common cause of pronunciation variation is reduction: words are often pronounced with fewer segments or syllables than their citation forms (for an introduction to the phenomenon of reduction, see Ernestus & Warner, 2011). For instance, the word *yesterday* can be pronounced in South-West American English as [jɛfeɪ]. Phonetic reduction leads to great variability in speech. Johnson (2004) showed that in a corpus of American English, no less than 20% of word tokens were pronounced with at least one segment missing, and 6% with at least one syllable missing. Similar results have been found for Dutch spontaneous speech (Schuppler, Ernestus, Scharenborg & Boves, 2011). In this study, we investigate whether highly reduced forms are prevalent in infant-directed speech, and we discuss the implications of this type of pronunciation variation for the building of a lexicon.

Despite the high frequency of occurrence of reduced pronunciation variants, these forms normally do not seem to hinder communication. Adult listeners appear to understand these variants without any problem, at least when they are heard in context (Ernestus, Baayen & Schreuder, 2002). This raises the question of when infants can start learning to process reduced variants. In this study we investigated whether words are also reduced in speech to 11- and 12-month-old infants, that is whether highly reduced pronunciation variants occur in infant-directed speech as well. We addressed this question by carrying out a detailed study of the reduced pronunciation (vowel

reduction and absence of vowels, consonants and syllables) of two word types, comparing spontaneous infant-directed speech, spontaneous adult-directed speech and adult-directed read speech.

These comparisons are relevant because knowledge about which pronunciation variants occur in infant-directed speech can provide insights into the development of phonological representations in the mental lexicon. A number of studies have shown that adult speakers and listeners are sensitive to the frequencies of different pronunciation variants, which makes it likely that they have stored these pronunciation variants in their mental lexicons. Ranbom and Connine (2007) conducted a series of perception experiments using American English words with an /nt/ cluster that can be pronounced as a nasal flap (e.g. *gentle* pronounced as [dʒɛrl]). They found a correlation between the frequency with which a given word occurs with the flap and the recognition times of that pronunciation variant. Bürki, Ernestus and Frauenfelder (2010) found similar results in a production experiment with French words that can be produced with or without a schwa in the initial syllable (e.g. *fenêtre* 'window' can be produced as [fənɛtʁ] or as [fnɛtʁ]). They showed that the relative frequencies of the pronunciation variants are correlated with response latencies: participants produced a variant faster the more frequent it is relative to the other variant. Finally, Pitt, Dilley and Tat (2011) investigated the correlation between the frequencies of different pronunciation variants of word-medial /t/ in American English ([t], [ʔ], [ɾ] and a deleted variant) and recognition in a lexical decision task. They found that the more commonly produced variants are also more easily recognized.

The assumption that pronunciation variants may be stored in the adult mental lexicon indicates a need for investigating the formation of these representations in language acquisition. An important question is, at what point in development are reduced pronunciation variants present in speech input

to children. If young infants' input mainly contains unreduced pronunciation variants for words that infants have not yet acquired, they will probably acquire the meaning and a phonological representation of the full pronunciation of a word before they form lexical representations for its reduced pronunciation variants. Once these variants become available in their linguistic input, they can link the representations for reduced variants to these word nodes. In contrast, if infant-directed speech contains unreduced as well as reduced pronunciation variants for highly frequent words which infants have not yet acquired, they can start acquiring reduced pronunciations early on in their linguistic development, and store these variants in their emerging mental lexicons. Thus, knowledge about the occurrence of phonetic reduction in infant-directed speech could provide important insights into how phonological representations in the mental lexicon develop.

If phonetic reduction occurs in infant-directed speech, another relevant question is whether the nature of this reduction resembles the nature of reduction in adult-directed speech. A number of studies (e.g. Havy, Bertoncini & Nazzi, 2011; Havy & Nazzi, 2009) suggest that consonant changes play a more critical role in word comprehension than vowel changes. Possibly, reduction affects mainly vowels in infant-directed speech whereas it also affects consonants in adult-directed speech.

A second reason to compare infant-directed speech with adult-directed speech with respect to phonetic reduction is that this comparison can lead to a better understanding of the nature and functions of infant-directed speech in language acquisition. Infant-directed speech is the primary input at early stages of language acquisition: this special speech register is used by caregivers cross-linguistically (Fernald et al., 1989), and many studies show that infants have a clear preference for infant-directed over adult-directed speech (e.g. Cooper & Aslin, 1990; Fernald, 1985; Werker, Pegg & McLeod, 1994). Researchers have suggested different functions for this speech register, namely getting and keeping the infant's attention, expression of positive affect, and clarification to facilitate

language acquisition (Soderstrom, 2007). A number of studies have found advantages of infant-directed over adult-directed speech in language learning: infant-directed speech has been shown to facilitate word segmentation (Thiessen, Hill & Saffran, 2005) and word recognition (Singh, Nestor, Parikh & Yull, 2009; Song, Demuth & Morgan, 2010). Although infants overhear other types of speech as well (e.g. speech from adults to other adults), these properties of infant-directed speech make it the most important source of information for infants.

Typical characteristics of infant-directed speech compared to adult-directed speech are utterances with fewer words, more pauses, an overall higher pitch, and exaggerated pitch contours (e.g. Fernald et al., 1989). On the segmental level, caregivers also use exaggeration. In infant-directed speech, speakers tend to produce more extreme variants of vowels, thereby increasing the acoustic distance between these vowels (e.g. Kuhl et al., 1997; but see Englund & Behne, 2006 and Dodane & Al Tamimi, 2007 for counter evidence). Furthermore, speakers may exaggerate the pronunciation of consonants in infant-directed speech (e.g. Bernstein Ratner, 1984; Cristià, 2010; but see Cristià & Seidl, 2013, for counterexamples), and have a greater tendency to produce canonical variants of consonants than in adult-directed speech, at least in read speech (Dilley, Millett, McAuley & Bergeson, 2013). In other words: in infant-directed speech, speakers appear to hyperarticulate segments to make them easier to distinguish, thereby accommodating to the needs of the listener (see Lindblom, 1990). These findings of exaggeration and careful pronunciation indicate that infant-directed speech contains fewer reduced pronunciation variants than casual adult-directed speech.

However, a number of studies also report reduced pronunciation in speech to infants and young children. Stoel-Gammon (1984) qualitatively described the acoustic properties of American English speech to 15- to 18-month-old children. She found that child-directed speech moves on a continuum from carefully articulated speech to casual speech, where the latter is characterized by processes of

reduction, assimilation and deletion. For instance, segments or sequences of segments were absent in frequently occurring phrases, such as *let's go* [sko] and *shall we* [fwi]. Stoel-Gammon does not report the frequency of occurrence of these heavily reduced forms nor is this a systematic study of their acoustic properties. It therefore remains unclear whether these reductions are exceptional in infant-directed speech.

Similar results of less careful articulation were reported for speech directed at somewhat older children. Shockey and Bond (1980) investigated processes of phonological reduction (substitution of glottal stop for word-final [t], loss of initial [ð], affrication, and deletion of [t] in the cluster [ts]) in speech to 2- to 4-year-old children and speech to adults. They found that in British English, mothers use single segment phonological reduction more when speaking to their child than when speaking to an experimenter. They propose that the use of phonological reduction may increase the tone of intimacy in a mother-child interaction. Similarly, Bard and Anderson (1983) studied the intelligibility of words extracted from British English speech directed at children between 1;10 and 3;0 years old and at an adult experimenter. Their results show that adult listeners less accurately identified words extracted from infant-directed speech than words extracted from adult-directed speech when these words were presented in isolation. In a follow-up study, Bard and Anderson (1994) demonstrated that when words directed at children aged 1;10 to 3;0 were more redundant because they were repeated or because they referred to a physically present entity, adults and children less often correctly identified them when presented out of context. This is in-line with findings for adult-directed speech, where it has been demonstrated that words tend to be more reduced the more predictable they are, given their frequency of occurrence but also given the context (e.g. Bell, Brenier, Gregory, Girand and Jurafsky, 2009; Pluymaekers, Ernestus and Baayen, 2005). Bard and Anderson's studies do not provide information about what acoustic properties caused these words to be less identifiable, and how and to what degree these words were reduced.

Several studies suggest that reduction in infant-directed speech is conditioned by the same factors as reduction in adult-directed speech. Pate and Goldwater (2011) also report predictability effects on word duration in American English infant-directed speech to 9- and 10-month-olds. Furthermore, Fisher and Tokura (1995) demonstrated that in English speech to 14-month-old infants, the difference between given and new words, measured by differences in duration of the stressed vowel, pitch range, maximum pitch and amplitude, from first to second mention of a word, was similar to adult-directed speech. Bortfeld and Morgan (2010), however, found that although there is a general trend to decrease stress from mention to mention, English-speaking mothers of 9- and 10-month-old infants alternated between stressed and unstressed words across mentions.

Together, these results strongly suggest that infant-directed speech contains reduced pronunciation variants. However, these reports on phonetic reduction in infant- and child-directed speech investigated either word duration or emphasis (Bortfeld & Morgan, 2010; Fisher & Tokura, 1995; Pate & Goldwater, 2011), mild reduction processes in speech directed at toddlers (Shockey & Bond, 1980), they did not provide quantitative data on reduction degree and frequency (Bard and Anderson, 1983, 1994; Stoel-Gammon, 1984) or focused on read instead of spontaneous speech (Dilley et al., 2013). Hence, they did not study the occurrence of pronunciation variants that are so reduced when compared to their full form that they could be classified as different words by infants. Moreover, several of these studies (Bard and Anderson, 1983, 1994; Shockey & Bond, 1980) focused on speech directed at children who have already developed lexical representations for a large number of words. Importantly, degree of reduction may change across age or developmental stage of the child, like many other variables in infant-directed speech (e.g. Stern, Spieker, Barnett & MacKain, 1983). These studies therefore leave open the question whether speech directed at young infants contains highly reduced pronunciation variants.

In order to start addressing this issue, we conducted a detailed study of phonetic reduction of two word types in Dutch speech addressed to 11- and 12-month-old infants. For these two word types, we investigated which segments and syllables may be absent or reduced, and the frequency of occurrence of reduced pronunciation variants. We focused on utterance-medial tokens of the Dutch words *allemaal* ('all', /ɑləmal/) and *helemaal* ('completely', /heləmal/). These words, a quantifier and an adverb, are neither function words, nor typical content words. These words are suitable as targets for our study for several reasons. First, they occur frequently in both infant-directed and adult-directed speech (the words *allemaal* and *helemaal* have frequencies of 1278 and 1059 per million in the Spoken Dutch Corpus; Oostdijk, 2000). Second, the infants in our speech corpus have probably not yet acquired semantic representations for these words, because the semantics of quantifiers and adverbs are typically not mastered until later in acquisition (by 27 months, *helemaal* is produced by only 25% of children, and *allemaal* by 44%; Schlichting & Iutje Spelberg, 2002). Third, previous research has shown that these words are likely to be highly reduced in conversational speech between adults. The word *allemaal* is often reduced to for instance [almal], [amal] or [amə], and *helemaal* to [hemal] or [hemə]. In the Ernestus Corpus of Spontaneous Dutch (Ernestus, 2000) *allemaal* is reduced to [aməl] in 62.7 % of tokens and *helemaal* to [heməl] in 20.0% of tokens (Schuppler et al., 2011). Finally, as these words are neither function words nor typical content words, are hardly ever the topic of a conversation and do not refer to a physical entity, their pronunciation is not likely to be affected by predictability or redundancy.

We compared infant-directed speech with both adult-directed informal speech and adult-directed read speech, as read speech generally involves more articulatory effort than spontaneous informal speech (Nakamura, Iwano & Furui, 2008; Strik, van Doremalen & Cucchiari, 2008; Warner & Tucker, 2011). Therefore, if infant-directed speech is characterized by hyperarticulation and clarity,

we expect it to be similar to read speech with respect to phonetic reduction. If, on the other hand, phonetic reduction is present in spontaneous, casual speech regardless of the addressee, we would expect to find a similar degree of reduction in infant-directed speech and spontaneous adult-directed speech.

We studied speech directed at 11- and 12-month-old infants. We chose this age group, because we are interested in whether and how frequently reduced pronunciation variants of a word are present in the input before the child acquires a semantic representation of that word. Our study also differs from earlier work in that it focused on Dutch, while all previous studies reporting reduction in infant-directed speech were based on either American or British English.

We used two different methodologies. In a corpus study, we investigated acoustic correlates of reduction, both durational and segmental. In a perception experiment, we presented the same tokens of *allemaal* and *helemaal* to adult listeners who rated them for reduction and provided phonetic transcriptions.

STUDY 1: ACOUSTIC ANALYSES

Method

Materials

We extracted utterances containing the Dutch words *allemaal* /aləmal/ ‘all’, and *helemaal* /heləmal/ ‘completely’ from a corpus of infant- and adult-directed Dutch (Johnson, Lahey, Ernestus, and Cutler, in prep.). This corpus contains video and audio recordings of interactions between an infant, at least one parent, and a secondary caregiver (the other parent, a grandparent, or a family friend). Caregivers were asked to teach their child four new words (a noun, a verb, an

adjective and a proper name; none of these words were target words in our study). Infants in two age groups participated: 11- and 12-month-olds (31 infants) and 18- and 19-month-olds (19 infants). During a short part of each recording session, an experimenter was also present. These experimenters were young student-assistants, and the conversations between them and the caregivers were quite informal. All infants visited the lab three times within one month. Each recording session lasted about 30 minutes. At the time of this study, only 23 recordings from the first recording session with the younger infants had been transcribed orthographically. Therefore, we only used materials from recordings made when the 11- and 12-month olds visited the lab for the first time. This represents approximately 13 hours of speech containing 57,649 word tokens and 3171 word types.

We first extracted all 230 tokens of the target words from these recordings. We then excluded those tokens that did not have a medial position in the utterance, since utterance-initial and utterance-final words are likely to be less reduced (e.g. Bell et al., 2003). We also excluded tokens produced in overlap with another speaker. In total, 52 tokens of infant-directed speech and 84 tokens of adult-directed speech were included in our analysis (see Table 1). These tokens were uttered by 42 different speakers. All of the speakers were from the Netherlands, 31 of them were aged between 20 and 44, 10 were older than 44, and the age of one speaker is unknown.

Furthermore, we included 80 tokens of *allemaal* and *helemaal* from adult-directed read speech. These materials were obtained from the subcorpus Library for the Blind of the Spoken Dutch Corpus (Oostdijk, 2000), which consists of parts of books read aloud in a lively manner. We restricted ourselves to utterance-medial tokens read by speakers from the Netherlands who were born after 1950, to match the group of speakers with the speakers of the spontaneous materials in

age and country of origin. These criteria yielded 97 tokens of *helemaal* and 70 of *allemaal* spoken by 47 different speakers. We randomly selected 40 tokens of each word type (see Table 1).

Both *allemaal* and *helemaal* can be pronounced with primary lexical stress on either the first or the final syllable. The majority of the tokens in our dataset had initial stress: only seven tokens of *allemaal* (6%) had final stress, and 24 tokens of *helemaal* (21%). Both *allemaal* and *helemaal* usually do not bear prosodic stress.

INSERT TABLE 1 AROUND HERE

The utterances from the corpus of infant- and adult-directed speech were classified as either infant- or adult-directed by the transcribers on the basis of the speech signal and the video recordings. In a control experiment, we verified whether the utterances containing our target words sounded like clear instances of the three registers. We asked 12 adult listeners to classify these utterances in one of the registers on the basis of the audio recordings. Since the recordings of read speech contained less background noise, the amount of background noise could form a cue to this register. In order to prevent this, we added background noise to most of the utterances of read speech. We therefore first classified the 136 stretches of spontaneous speech as containing no background noise or as containing background noise of one of four categories (white noise of a low, intermediate or high level, or white noise plus background noise, see Table 2), and calculated the proportion of utterances belonging to these categories. We then extracted noise representing each of these four categories from the corpus and mixed it with utterances of read speech in the same proportions.

INSERT TABLE 2 AROUND HERE

The 12 participants correctly identified utterances of read speech in 93.5% of the trials, utterances of spontaneous adult-directed speech in 83.0% of the trials, and utterances of infant-directed speech in 80.1% of the trials. Five utterances of adult-directed and two utterances of infant-directed speech were classified incorrectly by more than 75% of the participants. In order to investigate whether the high identification scores were caused by lexical differences between the speech registers instead of acoustic properties, we carried out a second control experiment, in which we asked 12 new participants to read written transcripts of the utterances and classify them as one of the three registers. This group of participants correctly identified read speech in 58.3% of the cases, adult-directed speech in 66.2% of the cases and infant-directed speech in 47.1% of the cases. The results of these control experiments indicate that the utterances were identifiable as instances of the three registers, and that this was not due to lexical differences alone.

Besides lexical differences, it is known that adult-directed spontaneous speech is produced faster than both infant-directed speech (e.g. Fernald & Simon, 1984) and adult-directed read speech. This was also true for our materials. Following Pluymaekers et al. (2005), we defined speech rate as the number of syllables per second in the utterance containing the token. The number of syllables was defined as the number of vowels in the phonetic transcription of the target word (see below) plus the number of vowels in the full pronunciations of the other words in the utterance. The mean speech rate for each register is given in Table 1. There were significant differences in speech rate between the conditions ($F(2, 213) = 13.37, p < 0.001$): post-hoc Tukey's tests revealed that adult-directed spontaneous speech had a higher speech rate than infant-directed speech and adult-directed read speech (both p 's < 0.001) and that infant-directed speech and adult-directed read speech did not differ significantly ($p > 0.1$).

When speech rate is high, words are more likely to be shorter and to be pronounced in a non-canonical way (e.g. Fosler-Lussier & Morgan, 1999). Therefore, if speech rate were the main predictor of degree of reduction, we would expect the adult-directed spontaneous speech in our sample to contain more reductions than the infant-directed and the adult-directed read speech. However, Van Son and Pols (1990, 1992) showed that a high speech rate does not always lead to more reduction. Therefore, although adult-directed speech has a higher speech rate than infant-directed speech, it is possible that these two registers are similar in their degree of reduction. If this is true, then this cannot be due to a difference in speech rate.

Measurements

The dependent variables in the acoustic study were the duration of the complete word, the number of syllables and the quality of the vowels in the initial and final syllables. A phonetician (the first author) extracted the target words from their utterances using the speech software package *Praat* (Boersma & Weenink, 2012). Boundaries between the target words and the neighboring words were determined on the basis of auditory information, visible changes in the waveforms and clear transitions in the spectrograms. Word duration was defined as the length of the extracted words in milliseconds. Because the durations did not show a normal distribution, we logarithmically transformed them, following Bell et al. (2009), amongst others.

The number of syllables and the quality of the vowels were obtained from phonetic transcriptions of the target words. The phonetician created transcriptions on the basis of the auditory information in the sound signal, supported by the visual information in the waveform and the spectrogram. A vowel was treated as reduced when the phonetician had transcribed it as a schwa, or when a tense vowel was transcribed as lax (in the word types *allemaal* and *helemaal* this concerns /a/ pronounced as /a/ and /e/ pronounced as /ɛ/ or /ɪ/). For verification, a second transcriber also phonetically

transcribed all tokens. All analyses reported below are based on the transcriptions of the first phonetician. Unless otherwise stated, the same effects were found for the data from the second transcriber.

Statistical modeling

We used linear regression, linear mixed effect regression (with Speaker as a random variable), and classification and regression tree analysis (CART) to evaluate the effects of speech register (infant-directed, adult-directed spontaneous and read speech) on the dependent variables.

Results

Word duration

Mean word durations were 271.4 ms for infant-directed speech (standard deviation 90.7 ms), 256.9 ms for adult-directed spontaneous speech (standard deviation 96.6 ms) and 362.3 ms for adult-directed read speech (standard deviation 84.7 ms). First, we fitted a linear mixed effect model to the dataset with log duration as the dependent variable, Speech Register (predictor of interest) and Word (*allemaal* or *helemaal*) as fixed factors, and Speaker as a random factor. This model revealed that the estimated variance for the factor Speaker was effectively zero, which is probably due to the large number of speakers who uttered only one token in the sample (48 of the 90 speakers). Because Speaker was the only random factor, we then used simple linear regression modeling, which revealed a significant effect of Speech Register ($F(2, 210) = 38.93, p < 0.001$). The tokens from read speech were significantly longer than the tokens from adult-directed spontaneous speech ($\beta = 0.2458, t = 3.867, p < 0.001$), as well as from infant-directed speech ($\beta = 0.2294, t = 3.147, p < 0.01$). There was no significant difference between infant-directed and adult-directed spontaneous speech ($\beta = -0.0163, t = -0.225, p > 0.1$).

Number of syllables

Figure 1 shows the proportions of transcriptions by number of syllables for each speech register². The pattern in Figure 1 is clear: tokens from infant-directed speech were similar to spontaneous adult-directed speech with respect to the number of syllables, and contained fewer syllables than adult-directed read speech. For instance, tokens of *allemaal* from both adult- and infant-directed spontaneous speech were pronounced as [ɑmə] and [ɑlma], and tokens of *helemaal* were pronounced as [hɪmə] and [emə].

INSERT FIGURE 1 AROUND HERE

This pattern was confirmed by statistical analysis. We analyzed the number of syllables in the transcriptions by means of a Classification and Regression Tree (CART, see e.g. Baayen, 2008). A classification tree splits the data into two parts several times on the basis of the values of the predictors. Each split results in two parts that are more homogeneous with respect to the dependent variable than the original part. If a split is statistically significant, the predictor can be assumed to statistically contribute to explaining the variance in the data. This type of analysis is non-parametric, and therefore it has the advantage that it does not assume normally distributed data. Our analysis resulted in a classification tree with one statistically significant split, indicating that adult-directed spontaneous speech and infant-directed speech differed from read speech with respect to the number of syllables. There was no other significant effect. The misclassification rate for this CART tree was 22.2%.

Quality of the vowel in the initial syllable

² There were a few tokens in adult-directed read speech which consisted of four syllables. In these tokens the final /l/ seems to form a syllable with a following word-initial schwa.

The initial vowel was present in all tokens. We analyzed the log odds ratio of the number of transcriptions with a full vowel to the number of transcriptions with a reduced vowel. Because nearly all tokens of *allemaal* contained a full initial vowel, we focused on the quality of the first vowel in tokens of *helemaal*. The number of word tokens in this analysis was 111. In infant-directed speech, the initial vowel was reduced in 63.0% of tokens, in adult-directed spontaneous speech in 77.3% of tokens, and in read speech in 42.5% of tokens. A logistic linear mixed effect model revealed a significant effect of Speech Register ($F(2, 108) = 4.6205, p < 0.05$). Adult-directed spontaneous speech contained more reduced vowels than read speech ($\beta = -1.5693, z = -3.055, p < 0.01$). There were no other significant effects.

Quality of the vowel in the final syllable

Then, we modeled the log odds ratio of the realization of the final vowel as full or reduced. We defined the final vowel as the vowel in the syllable starting with /m/ or its reduced pronunciation /w/. In seven cases, no syllable followed this segment, thus 209 observations were included in this analysis. The final vowel was reduced in 92.3% of tokens in infant-directed speech, 83.3% of tokens in adult-directed spontaneous speech and 65% of tokens in read speech. The logistic linear mixed effect model revealed a significant effect of Speech Register ($F(2,206) = 6.484, p < 0.01$). Read speech contained significantly fewer reduced final vowels than adult-directed spontaneous speech ($\beta = -0.8777, z = -2.328, p < 0.05$) and than infant-directed speech ($\beta = -1.9020, z = -3.298, p < 0.001$). There was no significant difference between infant-directed and adult-directed spontaneous speech ($\beta = 0.9816, z = 1.642, p > 0.1$).

Discussion

This acoustic analysis provides evidence for the presence of phonetic reduction in the Dutch words *allemaal* and *helemaal* directed at 11- and 12-month-old infants. For instance, tokens of *allemaal* in

infant-directed speech were pronounced as [ɑlma] and [ɑmə] and tokens of *helemaal* as [əmal], [hemɑ] and [emə]. Our measurements and transcriptions suggest that tokens of these words from infant-directed speech are more similar to tokens from spontaneous adult-directed speech than to tokens from adult-directed read speech with respect to word duration, number of syllables, and quality of the final vowel.

The question now arises whether tokens of *allemaal* and *helemaal* from spontaneous infant- and adult-directed speech are also perceived as having a similar degree of reduction. Therefore, we conducted a perception study with adult listeners. In this perception study, adult participants listened to all tokens, rated them for degree of reduction on a six-point scale, and provided a phonetic transcription of each token, using a transcription system based on Dutch orthography.

STUDY 2: PERCEPTION EXPERIMENT

Method

Participants

Thirty-six Dutch native speakers aged 18 to 26 (mean 20.8 years), who were not phonetically trained, participated in the experiment. They were all undergraduate students at Radboud University Nijmegen, reported no speech or hearing problems, and were paid for their participation.

Materials

Our material consisted of the 216 tokens of *allemaal* and *helemaal* analyzed in the acoustic study. These tokens were presented in isolation, except if the final /l/ was difficult to separate from the following word-initial vowel. In this case, this following vowel was included as well. Each

participant was randomly assigned to one of four experimental lists, each list presenting the experimental items in a different random order.

As described in the first study the recording quality of the tokens from read speech was considerably higher than that of the tokens from spontaneous speech, mostly due to the lack of background noise in the read speech. Since a difference in background noise could influence participants' ratings, we used the utterances mixed with (background) noise from the auditory control experiment.

We concatenated a square wave with a duration of 505 ms before each token (see e.g. Smits, Warner, McQueen and Cutler, 2003) to diminish the possibility that listeners would perceive the abrupt start of the speech stream as a click sound or a plosive (Pols and Schouten, 1978). The square wave consisted of an onset of five ms with gradually increasing amplitude, while the central part was 500 ms with a fixed amplitude of 52 dB. No square wave was added at the end of the words, to prevent a backward masking effect. The sound files were resampled to 44.1 kHz and their intensity was normalized to 70.0 dB.

The experimental lists were preceded by a practice session of five items. The practice items were utterance-final and utterance-initial tokens of *allemaal* and *helemaal* from the same corpora: two unreduced tokens from the Spoken Dutch Corpus, and two highly reduced tokens and one unreduced token from the corpus of infant- and adult-directed speech.

Procedure

Participants were tested individually in a sound attenuated experimental booth. They were seated in front of a computer screen. The auditory stimuli were presented with the software program E-prime

(Psychology Software Tools) through Sennheiser headphones. The participants were instructed to rate each stimulus for degree of reduction on a six-point scale, with 1 meaning "not reduced at all" and 6 meaning "highly reduced", using a computer keyboard. For a good visualization of the six-point scale, they were asked to use the alphanumeric part of the keyboard. In addition, the scale was presented on the computer screen.

After participants had provided a rating for a stimulus, the same stimulus was repeated another two times, and participants were asked to provide a phonetic transcription of the item. For this phonetic transcription, they were instructed to use an adapted version of Dutch spelling. In Dutch orthography, which is relatively shallow, the phonologically long vowels /a /, /e/, /o/ and /y/ are written with double letters ("aa", "ee", "oo", "uu"), except in open syllables, where they are spelled with a single letter (for instance in plurals: *raam* – *ramen*, [ra:m] – [ra:mə], "window" - "windows"). In our adapted spelling, phonologically long vowels were always spelled with double vowels, even in open syllables, and short vowels were spelled in capitals. Furthermore, we instructed participants to spell both schwa and /y/ as "U". A sheet with examples (not related to the target words) was present in the experimental booth. These examples can be found in the appendix. Participants could correct their transcriptions using the arrow keys or backspace.

There were two pauses in the experiment: one after 72 and one after 144 trials. The total duration of the experiment was approximately 45 minutes.

Results

Ratings

Two participants clearly used the scale in the opposite way as was indicated on the screen, as indicated by their ratings for the practice items and the correlation between the number of syllables

in their transcriptions and the rating scores. Therefore, we reversed the rating scores for these participants. Figure 2 shows the distribution of the rating scores for the three speech registers for all participants. The interrater agreement, as assessed by Kendall's coefficient of concordance, was substantial ($W = 0.801$).

INSERT FIGURE 2 AROUND HERE

Figure 2 clearly indicates that the rating scores for infant-directed speech resemble those for adult-directed spontaneous speech, and that read speech shows a completely different pattern. We confirmed this pattern in a statistical analysis: as the rating scores were not normally distributed, we analyzed them using a logistic linear mixed effect model with Speaker, Participant and Stimulus as random effects, in which we treated rating scores 1, 2 and 3 as Low Reduction and the scores 4, 5 and 6 as High Reduction. As expected given Figure 2, this model revealed a significant effect of Speech Register ($F(2,7773) = 87.784$, $p < 0.01$): adult-directed read speech was rated as less reduced than infant-directed speech ($\beta = -5.8549$, $z = -10.429$, $p < 0.001$) and than adult-directed spontaneous speech ($\beta = -5.5551$, $z = -11.111$, $p < 0.001$). There was no significant difference between infant-directed and adult-directed spontaneous speech ($p > 0.1$).³

Number of syllables

The number of syllables was defined as the number of vowels in the transcription. Figure 3 shows the proportions of transcriptions by number of syllables. The pattern in the figure is clear: the majority of transcriptions of read speech contained three syllables, whereas the transcriptions of infant-directed speech and spontaneous adult-directed speech mostly contained two syllables. Furthermore, the patterns for the latter two registers are highly similar. A statistical analysis

³ A logistic linear mixed effect regression of the dataset without the scores 3 and 4 and with 1 and 2 coded as Low Reduction and 5 and 6 as High Reduction yielded similar results.

confirms this. We treated transcriptions with one and two syllables as Reduced and three or more (see again footnote 2) syllables as Unreduced. We then analyzed the log odds ratio of the number of unreduced transcriptions to the number of reduced transcriptions with a linear mixed effect regression model with Speaker, Participant, and Stimulus as random effects and Speech Register and Word Type as fixed effects. As expected given Figure 3, this analysis revealed a significant effect of Speech Register ($F(2, 7773) = 83.657, p < 0.01$): tokens from read speech were transcribed with significantly more syllables than tokens from infant-directed speech ($\beta = 12.6374, z = 11.291, p < 0.001$) and than tokens from adult-directed speech ($\beta = 12.3017, z = 12.319, p < 0.001$). Importantly, there was no significant difference between infant-directed and adult-directed spontaneous speech ($p > 0.1$).

INSERT FIGURE 3 AROUND HERE

Quality of the vowel in the initial syllable

The initial vowel was transcribed as reduced in 29.0% of the trials for infant-directed speech, in 31.1% of the trials for adult-directed spontaneous speech and in 0.7% of the trials for read speech. We analyzed the log odds ratio of the number of transcriptions with a full vowel in the initial syllable to the number of transcriptions with a reduced vowel, using a logistic linear mixed effect model with Stimulus, Speaker and Participant as random effects. This model revealed a significant effect of Speech Register ($F(2, 7772) = 29.775, p < 0.001$): read speech contained fewer reduced vowels than adult-directed spontaneous speech ($\beta = -2.6656, z = -7.857, p < 0.001$) and than infant-directed speech ($\beta = -2.6155, z = -6.964, p < 0.001$). There was no significant difference between infant-directed and adult-directed spontaneous speech ($p > 0.1$). Furthermore, there was a significant effect of Word Type: tokens of *helemaal* were transcribed with more reduced vowels

than tokens of *allemaal* ($\beta = 2.6201$, $z = 9.485$, $p < 0.001$; 30.7% and 11.6% reduced vowels, respectively).

Quality of the vowel in the final syllable

Finally, we analyzed the log odds ratio of the number of transcriptions with a full vowel to the number of transcriptions with a reduced vowel in the final syllable. Final vowels were reduced according to 86.0% of the transcriptions of infant-directed speech, 83.2% of the transcriptions of adult-directed spontaneous speech and 57.3% of the transcriptions of read speech. A logistic linear mixed effect model revealed a significant effect of Speech Register ($F(2, 7773) = 24.79$, $p < 0.001$). Tokens from read speech were transcribed significantly less often with reduced vowels than tokens from adult-directed spontaneous speech ($\beta = -2.8406$, $z = -5.854$, $p < 0.001$) and than tokens from infant-directed speech ($\beta = -3.1841$, $z = -5.717$, $p < 0.001$). There was no significant difference between infant- and adult-directed spontaneous speech ($p > 0.1$).

Relating ratings and transcriptions

To investigate whether participants used the number of syllables and the quality of the vowels to determine the degree of reduction, we fitted a number of regression models, with the rating score as the dependent variable and these properties of the corresponding transcriptions as predictors. A separate model was run for each predictor as these properties are highly correlated. The analyses revealed that all predictors (the number of syllables, the quality of the first vowel, and the quality of the final vowel) significantly predict the rating scores. Given the Akaike Information Criterion, which determines the relative goodness of fit of a statistical model, the model with the number of syllables was the preferred model (see Table 3).

INSERT TABLE 3 AROUND HERE

GENERAL DISCUSSION

In this paper, we describe our investigation of phonetic reduction in infant-directed speech. In two studies, we analyzed in detail the pronunciation of the Dutch words *allemaal* ‘all’ /ɑləmal/ and *helemaal* ‘completely’ /heləmal/ in speech to 11- and 12-month-old infants, spontaneous adult-directed speech, and adult-directed read speech. Study 1 investigated the phonetic properties (duration, number of syllables and quality of the initial and final vowels) of 216 tokens of the target words extracted from corpora of spontaneous and read speech. We demonstrated that these words were approximately as short in infant-directed speech as in spontaneous adult-directed speech, but were longer in read speech. Furthermore, these words contained more syllables in read speech than in spontaneous infant- and adult-directed speech, while in read speech their vowels were also less often reduced. For instance, whereas the pronunciations [ɑmə] and [emə] occur frequently in infant- and adult-directed spontaneous speech, read speech contains many tokens of [ɑləmal] and [heləmal]. Our acoustic analysis thus shows that the reduction of the words *allemaal* and *helemaal* in infant-directed speech is more similar to adult-directed spontaneous speech than to read speech.

Study 2 investigated the perception of the same word tokens by adult listeners. Participants listened to all words, rated them for degree of reduction on a six-point scale, and provided a phonetic transcription of each word token using a transcription system based on Dutch orthography. The results showed that listeners perceive the same degree of reduction for tokens in infant-directed as in adult-directed spontaneous speech, but less reduction in read speech, as indicated by both reduction ratings and transcriptions.

These results reveal that infants are confronted with highly reduced pronunciation variants from early on in their lives, at least for some words. This result is striking, since prior studies have reported infant-directed speech to be carefully articulated in several respects. Contrary to findings of exaggerated pronunciation of vowels (e.g. Kuhl et al., 1997) and consonants (e.g. Bernstein Ratner, 1984; Cristià, 2010), we found that the majority of tokens of *allemaal* and *helemaal* from infant-directed speech were rated as reduced (see Figure 2) and contained one syllable less than the canonical pronunciation (see Figure 1 and Figure 3).

These results are also unexpected in the light of Lindblom's theory of Hyper and Hypo-articulation (Lindblom, 1990) if this theory had also been formulated to include child-directed speech. This theory assumes that speakers move along the dimension of hypospeech and hyperspeech and try to minimize their efforts, while maximizing comprehension. Compared to adults, infants have scarce linguistic experience, therefore, they might be assumed to need clearer input. Thus, when speakers communicate with an infant, this framework predicts that the speaker would use clear acoustic-phonetic patterns (hyperarticulation). Adults, on the other hand, have more experience with speech and are more able to predict the message. Speakers are thus predicted to use underarticulated hypospeech when speaking to adults, in situations where this is assumed to be enough for comprehension and successful communication. Yet our data show that adults may use as much reduction in infant-directed speech as in adult-directed speech, at least for some words.

Our findings extend earlier research on phonetic variability and reduced intelligibility in infant- and child-directed speech. Stoel-Gammon (1984) also found highly reduced pronunciation variants in infant-directed speech, but noted that very casual and very careful speech represent the two extremes of a continuum of styles, which constitute a minority of utterances. Our results suggest that highly reduced pronunciation variants may not be so exceptional in infant-directed speech, as

76% of the tokens of *allemaal* and *helemaal* examined in our study were produced with two instead of three syllables. Furthermore, our results show that highly reduced forms are present in speech to an even younger age group (11- and 12-month-olds) than in Stoel-Gammon's study.

Our results are in line with findings by Pate and Goldwater (2011), who found that the duration of words in infant-directed speech is affected by predictability effects which are, in many ways, similar to adult-directed speech. Our findings also show similarity between spontaneous adult-directed and infant-directed speech: in the extent to which the words *allemaal* and *helemaal* are reduced in duration.

Importantly, our data also show a high similarity between infant- and adult-directed speech in the type of reduction: segments, especially consonants, are equally often absent in both registers. This is in line with results by Dilley et al. (2013), who found similar distributions for non-canonical pronunciations in alveolar stops in infant- and adult-directed read speech. Bortfeld and Morgan (2010) speculate that variation in the input on dimensions that are not crucial for word identification (such as talker gender, pitch or emphatic stress) aids infants in distinguishing features that are lexically relevant from those that are not. The presence or absence of a segment, however, is a lexically relevant property of a word. From early in their development, infants are thus also confronted with pronunciation variation affecting lexically relevant properties of a word, which may complicate the acquisition of phonological representations.

As noted earlier, the utterances containing our target words in the three speech registers differed in speech rate (see Table 1) which is an important predictor for degree of reduction. Speech rate in the infant-directed speech in our sample was slower than in the spontaneous adult-directed speech, but it did not differ from the adult-directed read speech. Therefore, if speech rate were the main

predictor of degree of reduction, we would have expected the adult-directed spontaneous speech in our sample to contain more reductions than the infant-directed and the adult-directed read speech, which is the opposite of what we found. We therefore conclude that our results cannot be accounted for simply by differences in speech rate between infant-directed, spontaneous adult-directed and read speech.

In this study, we only investigated two words which are similar in many respects: number of phonemes, number of syllables, and stress pattern. Furthermore, both words often appear in non-accented positions in the sentence, and are therefore prone to reduction in adult-directed speech (e.g. Bell et al., 2003). Finally, the two words are also similar in that they are neither typical content words nor typical function words. As a consequence, our results cannot easily be generalized to other words. Nevertheless, initial evidence suggests that canonical content words may also be produced in variable ways (Bortfeld & Morgan, 2010; Fisher & Tokura, 1995). Future research should directly compare reduction in infant- and adult-directed speech on the basis of a greater number of word types.

If infant-directed speech is indeed comparable to adult-directed speech with respect to reduction for many word types, this has consequences for language acquisition. Great variability in pronunciation arising from phonetic reduction poses challenges to the language learning infant, namely the building of phonological representations of words despite the different pronunciation variants in the input, and the recognition of different pronunciation variants as instances of the same word.

Consider the word *allemaal*, a highly frequent word whose meaning 12-month-old infants most probably do not yet know, but which occurs frequently enough for infants to store. Our results show that infants' input contains both unreduced pronunciation variants and reduced pronunciation

variants for this word (e.g. [ɔləmal] and [ɔmə]). The reduced pronunciation variants occur so frequently, that infants probably store these variants.

Support for the assumption that infants store frequent reduced pronunciation variants comes from several lines of research. First of all, there is evidence that adults store representations of reduced pronunciation variants in their mental lexicons (Bürki et al., 2010; Pitt et al., 2011; Ranbom and Connine, 2007). Given the high frequency of reduced pronunciation variants in our data, it could well be the case that these representations are acquired in infancy. A second argument for the hypothesis that infants store reduced pronunciation variants of *allemaal* and *helemaal* is that several studies have shown that 10- to 11-month-old infants' phonological representations are detailed at the segmental level (Johnson, 2005; Shafer, Shucard, Shucard and Gerken, 1998; Shi, Werker and Cutler, 2006; Swingley, 2005) and the reduced pronunciation variants often differ from the citation forms at the segmental level. A third reason to assume early storage of reduced pronunciation variants is formed by a recent study of French child speech (Dye, 2011). This study suggests that the different realizations for auxiliaries by children aged between 23 and 37 months may be related to the reduction of these elements in adult-directed speech. The children in this study are, however, much older than the infants in our study. Experimental work with infant listeners is needed to shed more light on how young infants exactly deal with variation caused by phonetic reduction.

There are several possibilities for the storage of reduced pronunciation variants by infants. A first possibility is that infants' early representations of words are holistic or segmentally underspecified (see Hallé and De Boysson-Bardies, 1996; but see the literature mentioned in the previous paragraph). For instance, a representation could consist of a word's syllabic structure and consonantal and vocalic slots which are underspecified for content. Our data suggest that phonetic reduction causes so much pronunciation variation that it is unlikely that infants can map the reduced

and unreduced forms to a singular holistic representation, since the variants differ in their number of syllables as well as in the numbers of vowels and consonants. A second possibility, which seems more likely given our data, is based on the assumption that highly reduced pronunciation variants differ so much from unreduced pronunciation variants that infants may have difficulty inferring that these variants represent the same word. Furthermore, the unreduced and reduced variants may slightly differ in their pragmatic function (e.g. Plug, 2005). These acoustic and pragmatic differences between reduced and unreduced pronunciation variants may lead infants to store these variants as separate entries in their mental lexicons. At a later stage they then have to merge these representations to a single word node which carries the semantic and syntactic information of the word. In contrast, the third possibility is based on our finding that infants' input contains not only highly reduced pronunciation variants (e.g. [ɑmə]) and unreduced variants, but also variants with a degree of reduction that lies in-between these variants. The variants present in infant-directed speech thus form a continuum from unreduced to reduced variants. The infant could pick up on this and therefore store the different pronunciation variants as being variants of the same word from early on in linguistic development.

In conclusion, this paper has shown that phonetic reduction is present in the Dutch word types *allemaal* and *helemaal* in speech directed at 11- and 12-month-old infants, and that their degree of reduction is similar to spontaneous adult-directed speech. These findings extend the research on reduction in conversational speech between adults to the register of infant-directed speech. They show that infants are confronted with reduced word forms from early on in language acquisition. These findings suggest that representations for reduced pronunciation variants (for at least some words) can be acquired early on in linguistic development.

ACKNOWLEDGMENTS

This research was funded by a European Young Investigator Award to the second author, and an NWO Spinoza Grant entitled ‘Native and Non-native Listening’, awarded to Anne Cutler. We would like to thank Anne Cutler for her valuable comments on this paper, and Jens Frielink for being the second transcriber.

REFERENCES

- Baayen, R.H. (2008). *Analyzing linguistic data: a practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Bard, E.G., & Anderson, A.H. (1983). The unintelligibility of speech to children. *Journal of Child Language*, 10, 265-292.
- Bard, E.G., & Anderson, A.H. (1994). The unintelligibility of speech to children: effects of referent availability. *Journal of Child Language*, 21, 623-648.
- Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America*, 113, 1001–1024.
- Bell, A., Brenier, J., Gregory, M., Girand, C., Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, 60, 92-111.
- Bernstein Ratner, N. (1984). Patterns of vowel modification in mother-child speech. *Journal of Child Language*, 11, 557-578.
- Boersma, P. & Weenink, D. (2012). Praat: doing phonetics by computer [Computer program]. Version 5.3.32, retrieved 17 October 2012 from <http://www.praat.org/>
- Bortfeld, H., & Morgan, J. L. (2010). Is early word-form processing stress-full? How natural variability supports recognition. *Cognitive Psychology*, 60, 241-266.
- Bürki, A., Ernestus, M., & Frauenfelder, U. H. (2010). Is there only one "fenêtre" in the production lexicon? On-line evidence on the nature of phonological representations of pronunciation variants for French schwa words. *Journal of Memory and Language*, 62, 421-437.
- Cooper, R.P., & Aslin, R.N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, 61(5), 1584-1595.
- Cristià, A. (2010). Phonetic enhancement of sibilants in infant-directed speech. *Journal of the Acoustical Society of America*, 128(1), 424-434.

- Cristia, A. & Seidl, A., (2013). The hyperarticulation hypothesis of infant-directed speech. *Journal of Child Language*, Available on CJO doi:10.1017/S0305000912000669.
- Dilley, L.C., Millett, A.L., McAuley, J.D., & Bergeson, T.R. (2013). Phonetic variation in consonants in infant-directed and adult-directed speech: the case of regressive place assimilation in word-final alveolar stops. *Journal of Child Language*. Available on CJO 2013 doi:10.1017/S0305000912000670
- Dodane C. & Al-Tamimi J. (2007) An acoustic comparison of vowel systems in adult-directed-speech and child-directed-speech: Evidence from French, English & Japanese. In: In J. Trouvain & W. J. Barry (eds.), *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS 2007)*, Saarbrücken, Germany, 1573-1576.
- Dye, C.D. (2011). Reduced auxiliaries in early child language: Converging observational and experimental evidence from French. *Journal of Linguistics*, 47(2), 301-339.
- Englund, K. & Behne, D. (2006). Changes in infant directed speech in the first six months. *Journal of Infant and Child Development*, 15(2), 139-160.
- Ernestus, M. (2000). *Voice Assimilation and Segment Reduction in Casual Dutch: A Corpus-Based Study of the Phonology-Phonetics Interface*. Utrecht: LOT.
- Ernestus, M., Baayen, R.H., & Schreuder, R. (2002). The recognition of reduced word forms. *Brain and Language*, 81(1-3), 162-173.
- Ernestus, M., & Warner, N. (2011). An introduction to reduced pronunciation variants. *Journal of Phonetics*, 39, 253–260.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, 8, 181-195.
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20, 104-113.
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16(3), 477–501.
- Fisher, C., & Tokura, H. (1995). The given/new contract in speech to infants. *Journal of Memory and Language*, 34, 287-310.
- Fosler-Lussier, E., & Morgan, N. (1999). Effects of speaking rate and word frequency on pronunciations in conversational speech. *Speech Communication*, 29, 137-158.
- Hallé, P.A., & de Boysson-Bardies, B. (1996). The format of representation of recognized words in infants' early receptive lexicon. *Infant Behavior and Development*, 19, 463–481.
- Havy, M., & Nazzi, T. (2009). Better processing of consonantal over vocalic information in word learning at 16 months of age. *Infancy*, 14(4), 439 – 456.

- Havy, M., Bertoncini, J., & Nazzi, T. (2011). Word learning and phonetic processing in preschool-age children. *Journal of Experimental Child Psychology*, 108, 25–43
- Johnson, E.K. (2005). English-learning infants' representations of word forms with iambic stress. *Infancy*, 7(1), 95-105.
- Johnson, E.K., Lahey, M., Ernestus, M., & Cutler, A. (in preparation). The input for word learning: A new corpus of speech to infant listeners.
- Johnson, K. (2004). Massive reduction in conversational American English. Spontaneous speech: data and analysis. In K. Yoneyama & K. Maekawa (eds.), *Spontaneous speech: Data and analysis. Proceedings of the 1st session of the 10th international symposium* (pp. 29-54). Tokyo, Japan: The National International Institute for Japanese Language.
- Kuhl, P.K., Andruski, J.E., Chistovich, I.A., Chistovich, L.A., Kozhevnikova, E.V., Ryskina, V.L., Stolyarova, E.I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277, 684-686.
- Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In W.J. Hardcastle, & A. Marchal (eds.), *Speech Production and Speech Modeling* (pp. 403-440). Dordrecht: Kluwer.
- Nakamura, M., Iwano, K., & Furui, S. (2008). Differences between acoustic characteristics of spontaneous and read speech and their effects on speech recognition performance. *Computer Speech and Language*, 22, 171-184.
- Oostdijk, N. (2000). The Spoken Dutch Corpus Project. *The ELRA Newsletter*, 5(2), 4-8.
- Pate, J., & Goldwater, S. (2011). Predictability effects in adult-directed and infant-directed speech: Does the listener matter? In L. Carlson, C. Holscher & T. Shipley (eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society, 1569-1574.
- Pitt, M.A., Dilley, L., & Tat, M. (2011). Exploring the role of exposure frequency in recognizing pronunciation variants. *Journal of Phonetics*, 39, 304-311.
- Plug, L. (2005) From words to actions: The phonetics of eigenlijk in two pragmatic contexts. *Phonetica*, 62, 131-145.
- Pluymaekers, M., Ernestus, M., & Baayen, R.H. (2005). Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America*, 118(4), 2561-2569.
- Pols, L.C.W., & Schouten, M.E.H. (1978). Identification of deleted consonants. *Journal of the Acoustical Society of America*, 64, 1333–1337.
- Ranbom, L. J., & Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language*, 57, 273–298.
- Schuppler, B., Ernestus, M., Scharenborg, O., & Boves, L. (2011). Acoustic reduction in conversational Dutch: A quantitative analysis based on automatically generated segmental

- transcriptions. *Journal of Phonetics*, 39(1), 96-109.
- Shafer, V.L., Shucard, D.W., Shucard, J.L., & Gerken, L.A. (1998). An electrophysiological study of infants' sensitivity to the sound patterns of English speech. *Journal of Speech, Language, and Hearing Research*, 41, 874-886.
- Schlichting, J.E.P.T. & Iutje Spelberg, H.C. (2002). *Lexilijst Nederlands. Een instrument om de taalontwikkeling te onderzoeken bij Nederlandstalige kinderen van 15-27 maanden in het kader van vroegtijdige onderkenning*. [Lexilijst Dutch. A tool to study language development in Dutch children aged 15-27 months in the framework of early detection]. Lisse: Swets Test Publishers.
- Shi, R., Werker, J.F., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy*, 10(2), 187-198.
- Shockey, L., & Bond, Z.S. (1980). Phonological processes in speech addressed to children. *Phonetica*, 37, 267-274.
- Singh, L., Nestor, S., Parikh, C., & Yull, A. (2009). Influences of infant-directed speech on early word recognition. *Infancy*, 14(6), 654-666.
- Smits, R., Warner, N., McQueen, J., & Cutler, A. (2003). Unfolding of phonetic information over time: A database of Dutch diphone perception. *Journal of the Acoustical Society of America*, 113(1), 563-574.
- Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, 27, 501-532.
- Song, J.Y, Demuth, K. & Morgan, J. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *Journal of the Acoustical Society of America*, 128(1), 389-400.
- Stern, D.M., Spieker, S., Barnett, R.K., & MacKain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language*, 10, 1-15.
- Stoel-Gammon, C. (1984). Phonological variability in mother-child speech. *Phonetica*, 48, 7.
- Strik, H., van Doremalen, J., & Cucchiaroni, C. (2008). Pronunciation reduction: how it relates to speech style, gender, and age. In J. Fletcher, D. Loakes, R. Gocke, D. Burnham, & M. Wagner (eds.) *Proceedings of Interspeech 2008* (pp. 1477-1480). Brisbane: ISCA.
- Swingle, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, 8(5), 432-443.
- Thiessen, E.D., Hill, E.A., & Saffran, J.R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, 7, 53-71.
- Van Son, R.J.J.H. & Pols, L.C.W. (1990). Formant frequencies of Dutch vowels in a text, read at normal and fast rate. *Journal of the Acoustical Society of America*, 88, 1683-1693.

Van Son, R.J.J.H. & Pols, L.C.W. (1992). Formant movements of Dutch vowels in a text, read at normal and fast rate. *Journal of the Acoustical Society of America*, 92, 121-127.

Warner, N. & Tucker, B.V. (2011). Phonetic variability of stops and flaps in spontaneous and careful speech. *Journal of the Acoustical Society of America*, 130(3), 1606-1617.

Werker, J.F., Pegg, J.E., & McLeod, P.J. (1994). A cross-language investigation of infant preference for infant-directed communication. *Infant Behavior and Development*, 17, 323-333.

Appendix A. List of words on the example sheet

Table 4 presents the words that were present on the example sheet in their conventional spelling and in the adapted spelling system. The canonical pronunciations (IPA) were not on the example sheet, but are provided here for convenience.

TABLE 4

Word	Adapted spelling	Pronunciation
bel	bEl	bɛl
beeld	beeld	bɛ:lɪ
bal	bAl	bɑl
vaas	vaas	vɑ:s
grot	grOt	χrɔt
groot	groot	χro:t
pit	pIt	pɪt
rug	rUg	rʏχ
kale	kaalU	ka:lə

TABLE 1. Number of tokens and mean speech rate of these tokens' utterances (in syllables per second) by speech register.

	<i>Allemaal</i>	<i>Helemaal</i>	Mean speech rate (standard deviation)
Adult-directed spontaneous speech	40	44	6.37 (1.55)
Infant-directed spontaneous speech	25	27	5.65 (1.20)
Adult-directed read speech	40	40	5.38 (0.84)

TABLE 2. Number of utterances in read speech by noise category.

	Number of utterances
White noise, low level	19
White noise, intermediate level	18
White noise, high level	9
White noise plus background noise	15
No noise added	19
Total	80

TABLE 3. Akaike information criterion for each of the models.

Predictor	AIC
Number of syllables	7074
Quality of the first vowel	9841
Quality of the final vowel	9212

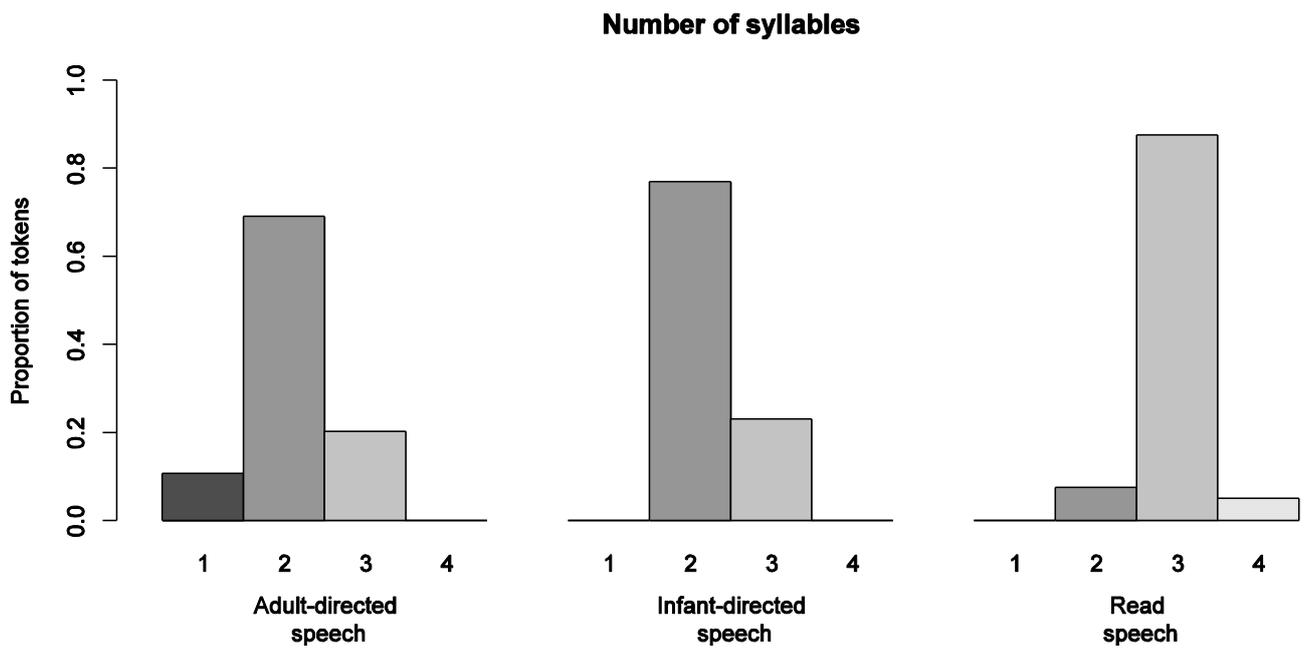


FIGURE 1. Number of syllables for the three speech registers in the phonetician's transcriptions.

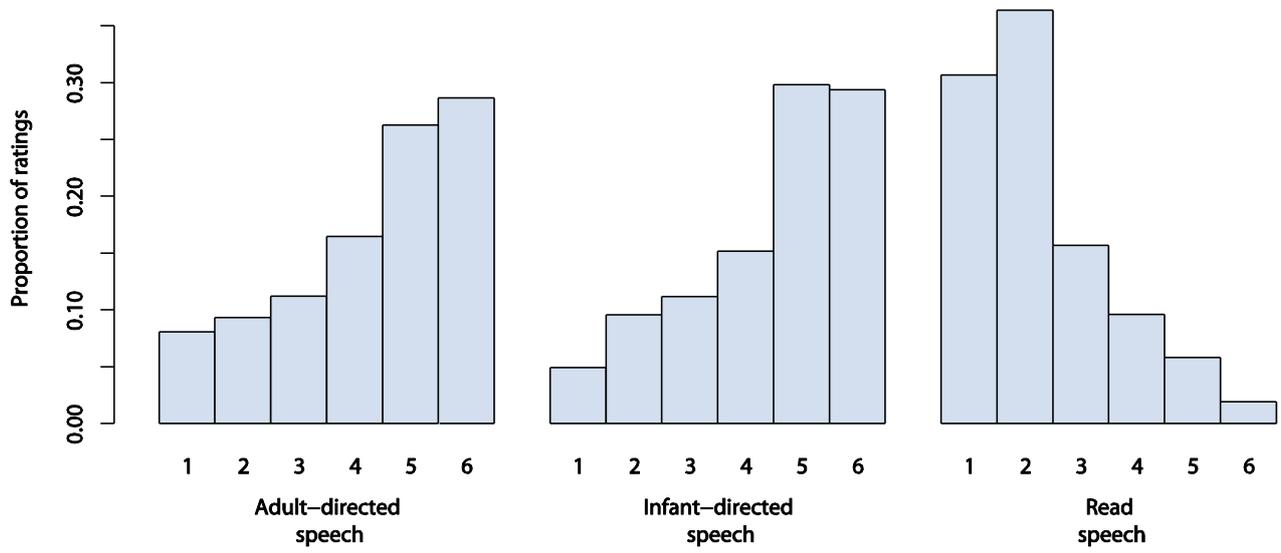


FIGURE 2. Proportions of rating scores for the three speech registers. (1 = “not reduced at all”, 6 = “highly reduced”).

Number of syllables

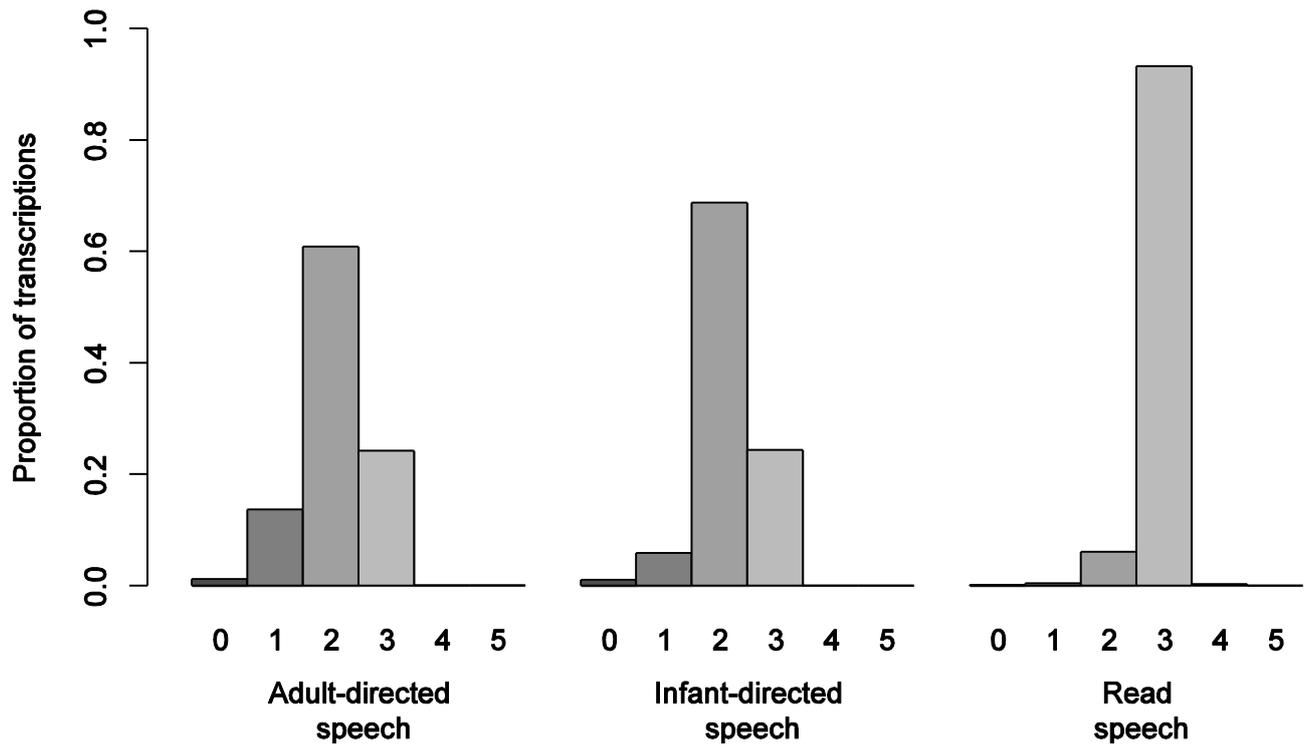


FIGURE 3. Number of syllables for the three speech registers in the participants' transcriptions.