PATERNAL INFANT-DIRECTED SPEECH:

An acoustic study on the vowels of Dutch parents

MARLOES ROOSINGH

Student number 10455450

BA thesis Linguistics

Supervised by Titia Benders

University of Amsterdam, 27th of June 2022

ABSTRACT

Fathers, similar to mothers, engage in infant-directed speech (IDS) which is characteristically different from adult-directed speech (ADS). The role of fathers' speech in children's language development, and many of the specific features of their IDS, are currently under-researched. Aiming to extend our understanding of the characteristics of paternal IDS, the present study investigates how Dutch fathers articulate their vowels when speaking to 15-month-old infants. The corner vowels of 5 fathers and 5 mothers in IDS and in ADS were analyzed and compared. The results provide no evidence that Dutch fathers amend their vowel space in IDS, as compared to ADS, and as compared to mothers.

The results and method of this study were reflected on and considered in light of previous research, highlighting several caveats of both. Recommendations for future research include consideration of the background of participants, and increased transparency on method of vowel space measurement.

TABLE OF CONTENTS

1. Introduction

Infant-directed speech is acoustically distinguishable from adult-directed speech by modifications in F0 (fundamental frequency), F0 variability, vowel space area, articulation rate, and vowel duration (as reviewed in Cox et al. 2022). Children have a preference for infant-directed speech over adult-directed speech (Fernald 1985; The ManyBabies Consortium 2020; Byers-Heinlein et al. 2021). The acoustic profile of infant-directed speech is widely assumed to be conducive to child language acquisition, its various properties being thought to draw the child's attention and facilitate the learning of certain language features (Saint-Georges et al. 2013).

While infant-directed speech is clearly different from adult-directed speech at the prosodic level, its expression at the level of segments is less straightforwardly classified. Where the higher fundamental frequency and enhanced intonation can generally be regarded as exaggerated or *hyper-articulated* relative to adult-directed speech, there is no consensus on the articulation of single phonetic units such as vowels (Cristia 2013), and recent studies suggest that in some cases, they may be *hypo-articulated* instead (Cristia & Seidl 2014; Englund 2018).

The vowel space as the surface area between the outermost vowels in a F1xF2 plane is often used to determine whether a speaker is *hyper-articulating*; enhancing the contrasts between sounds and producing clearer speech, measurable as a larger vowel space, or *hypo-articulating*; producing diminished contrasts and speaking less clearly, measurable as a smaller vowel space.

1.1 Vowel space across languages

Hyper-articulation in infant-directed speech has previously been claimed to be universal (Kuhl et al. 1997; Uther, Knoll & Burnham 2007), and has been summarized into the *hyper-articulation hypothesis* by Cristia & Seidl (2014): supposedly, hyper-articulation promotes language acquisition by presenting enhanced contrasts between phonemic categories. The vowel space is indeed reported to be larger in infant-directed speech in a number of languages, such as American English, Russian, and Swedish (Kuhl et al. 1997), Australian English (Xu et al. 2013), Hungarian (Gergely et al. 2017), Mandarin Chinese (Liu, Kuhl & Tsao 2003), and Spanish and Basque (Kalashnikova & Carreiras 2021). Upon casting a wider net, however, this feature is not necessarily found everywhere. In Norwegian and in Dutch, the infant-directed vowel space has actually been found to be smaller than in adult-directed speech (Englund & Behne 2006; Benders 2013).

1.2 Vowel space in infant-directed speech

Adults appear to purposefully adjust their vowel space in infant-directed speech, rather than solely as a characteristic of the register used with something perceived to be ‘cute’ or small. For example, dog-directed speech shares many features with infant-directed speech, but a Hungarian study shows an amended vowel space is not one of them (Gergely et al. 2017), and it has been shown that mothers who used a larger vowel space with their children in Australian English did not do the same when talking to pets (Xu et al. 2013). In comparing the vowel space between speech directed at adults, at dogs, at parrots, and at infants, the vowel space increased in that order, indicative of a relationship between the expected potential linguistic competency of the latter three groups and the vowel space used by the mothers. Additionally, a study on a mother with twins, one hearing-impaired and one hearing, showed she amended her vowel space to the child; she did not use a larger than typical vowel space when addressing her hearing-impaired infant, while she did use a larger vowel space with the hearing infant (Lam & Kitamura 2010). This last example is an anecdotal illustration of the deliberateness, in that the mother did not hyper-articulate with the child that was not expected to hear her properly. These findings suggest there is a communicative goal to using an amended vowel space in infant-directed speech.

1.3 Hypo-articulation

In recent years it has been argued by some that hypo-articulation may actually commonly occur in infant-directed speech (Englund 2018), even in languages that report a trend of hyper-articulation. Plus, where hyper-articulation is believed to aid child language acquisition (see the hyper-articulation hypothesis in §1.1), it's not unequivocally proven to do so. A Swedish study reports the relationship between parent hyper-articulation and child language production to be negative; parent vowel hyper-articulation led to diminished complexity in the following infant vocalization. The authors suggest this effect stems from the abundance of phonetic information that hyper-articulation provides, which may hinder its reception rather than facilitate it (Marklund, Marklund & Gustavsson 2021).

A more differentiated position is the perspective offered by Cristia & Seidl (2014), who found evidence of hyper-articulation in American English in the corner vowels, but not across all vowels: for example, the [i–ɪ] contrast was hypo-articulated. Englund (2018) also advocates for considering the vowel space beyond the corner vowels, especially in languages where hyper-articulation seems to be the trend for infant-directed speech. These observations and caveats lead to a careful conclusion that hypo-articulation may be more prevalent than previously thought, and possibly even serves its own didactic function in infant-directed speech.

1.4 Paternal infant-directed speech

Out of the studies on vowel space mentioned in §1.2, 6 out of 7 studies were carried out with mothers and their infants, with the single remaining one also including fathers. This ratio approximates a larger trend in research on infant-directed speech; Saint-Georges et al. (2013) report that as few as 7 out of 114 studies included paternal infant-directed speech. While tentative steps have been made over the past few decades (the review by Saint-Georges et al. considers studies from 1966 onwards, the first one to consider fathers is from 1982), the available data on paternal infant-directed speech is still scarce, leaving the topic open to further investigation.

One area in which studies do show differences between fathers and mothers is the quantity of speech that infants are exposed to. And interestingly, there seems to be a correlation between paternal speech input and child language development in several studies where the child gets less input from the father than from the mother. Shapiro, Hippe & Ramírez (2021) found that even when fathers contributed a third of the total amount of speech heard, it was a better indicator of the child's language production than the mother's contribution, and Pancsofar & Vernon-Feagans (2006) found that while the total amount of speech input from fathers was lower on average, fathers' vocabulary use was positively correlated with child language development whereas that of mothers was not.

In light of this correlation between the paternal speech input and child language development, the question arises of why paternal infant-directed speech may have a different impact, and whether it has different qualities than maternal speech.

1.5 Aim of the current study

The current study is aimed at shedding more light on the physical characteristics of paternal infant-directed speech through a closer look at its phonetic properties as compared to maternal infant-directed speech and adult-directed speech in general. Collecting this information is vital to the question of why paternal infant-directed speech may have a specific impact. After all, in order to be able to draw conclusions on why a phenomenon has a certain impact, we need a clear overview of the properties of said phenomenon. Additionally, if further examples of a consistently smaller vowel space in infant-directed speech are found, that would provide support for more extensive inquiries into hypo-articulation in infant-directed speech.

The vowels /i/, /u/, /a/, and /ɑ/ mark the corners of the Dutch vowel space, as illustrated in Figure 1.

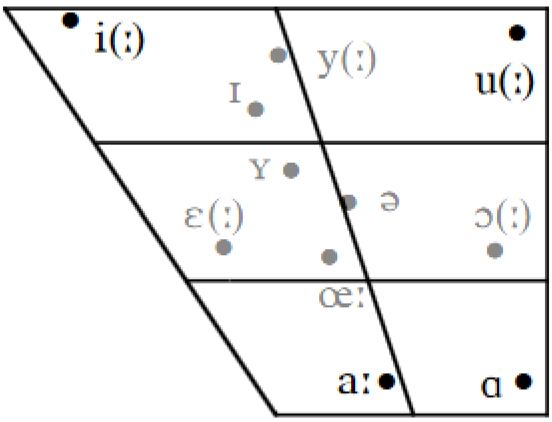


Figure 1: the relative position of Dutch monophthongs in the vowel space, adapted from Gussenhoven (1992).

To determine how Dutch fathers use this vowel space, we measure the first and second formant (F1 and F2) in the corner vowels from a pre-existing corpus of recordings of parents' utterances directed at their infants, and utterances directed at adults. We then compare the paternal infant-directed vowel space to their adult-directed vowel space, as well as to the maternal vowel spaces in both registers.

As Dutch mothers were found to hypo-articulate in infant-directed speech (Benders 2013), we hypothesize that Dutch fathers do so as well. Additionally, because men are found to hypo-articulate compared to women (Jacobi 2009; Escudero et al. 2009; Kempe, Puts & Cárdenas 2013; Weirich & Simpson 2018) we hypothesize that fathers use a smaller vowel space than mothers in both registers. We do not expect to find an interaction effect between register and gender on the size of the vowel space.

2. METHOD

2.1 DESIGN

The current study makes use of an existing corpus (Benders, StGeorge & Fletcher 2021) which contains speech from each of the conditions that are relevant to our question: fathers addressing an adult, fathers addressing an infant, mothers addressing an adult, and mothers addressing an infant. The recordings were originally collected in the context of a speech perception experiment, and have been analyzed previously in the context of pitch variability in paternal infant-directed speech. More information on this corpus can be found in Benders, StGeorge & Fletcher (2021).

The vowels /i/, /u/, /a/, and /ɑ/ mark the corners of the Dutch vowel space, as previously illustrated in Figure 1. The surface area between these corner vowels is the dependent variable in a 2x2 design with the gender of the speaker (male or female) as the between-subject independent variable, and the register (infant-directed or adult-directed) as the within-subject independent variable.

2.2 PARTICIPANTS

The participants in this study are all monolingually raised native speakers of Dutch who volunteered their time for a speech perception experiment unrelated to infant-directed speech. The complete dataset contains recordings of 28 caregivers each accompanied to the recording session by a single child. For this study, 5 father-infant pairs (3 daughters; 2 sons), and 5 mother-infant pairs (2 daughters; 3 sons) were selected to constitute a sample primarily balanced for parent gender. The pairs were chosen to have the ages of the infants as close together as possible; all children were approximately 15 months old (mean age: 462.9 days, range: 449–469 days) at the time of recording. The infants would ideally also be split evenly on gender, but have been selected to be as balanced as possible within the limitations of the chosen age group within the corpus. Infant age has been given preference over infant gender, because the vowel space in infant-directed speech has been reported to change with the age of the child (Cristia & Seidl 2014; Gergely et al. 2017), but in other cases also shown to remain steady (Kalashnikova & Burnham 2018; Benders 2013). To completely rule out any differences that may be accounted for by infant age, the age variation has been kept as small as possible for the current study.

The parents informally reported they were the primary caregiver for at least one day a week, excluding weekends. Due to the nature of the experiment the participants signed up for, no additional information about anything else was collected.

2.3 STIMULI

The items used in the recording phase provided each interaction with several words containing the target vowels /i/, /u/, /a/, and /ɑ/, as specified in Table 1. The words in the dataset corresponding to these stimuli are a mix of words between one and four syllables. The elicited recordings contain subsets of these words, as well as plurals and compounds containing them, and derived diminutives.

Table 1: words corresponding to the stimuli, containing the target vowels, each given in IPA transcription, with Dutch spelling underneath, and their English translation in italics.

Vowel	Items						
/i/	/fits/	/xitər/	/spixəl/	/sinasɑpəl/			
	fiets	gieter	spiegel	sinaasappel			
	bike	*watering can*	*mirror*	*orange*			
/u/	/buk/	/vut/	/ku/	/pus/			
	boek	voet	koe	poes			
	book	*foot*	*cow*	*cat*			
/a/	/aːp/	/kaːs/	/sxaːp/	/taːfəl/	/raːtəl/		
	aap	kaas	schaap	tafel	ratel		
	monkey	*cheese*	*sheep*	*table*	*rattle*		
/ɑ/	/ɑpəl/	/bɑt/	/jɑs/	/kɑt/	/slɑk/	/tɑs/	/slɑp/
	appel	bad	jas	kat	slak	tas	slab
	apple	*bath*	*jacket*	*cat*	*snail*	*bag*	*bib*

2.4 EQUIPMENT AND PROCEDURE

All recordings were made in a soundproof studio, with an omni-directional Samson QV microphone. The microphone was attached to the head of the parent and had a long cord to the remaining equipment for ease of movement. The sound was sampled at 44,000 Hz in the program Enosoft DV Processor.

The recording phase was prefaced by a short conversation between the experimenter and the parent, to allow them to get used to their surroundings and the studio setting. The parent and infant were seated on a blanket on the floor. In the recording phase, the parent and infant were left alone for approximately 10 minutes. The interactions were guided towards elicitation of the target words by the unpacking of three bags containing items, and the naming of these items (specified in Table 1).

Upon re-entry, the experimenter asked the parent about the session, to elicit adult-directed speech about the same items.

2.5 Annotation

The recordings were divided into utterances with a script (de Jong & Wempe 2009) in the speech analysis program Praat (Boersma & Weenink 2022). The resulting utterance boundaries were evaluated and adjusted where necessary. Each utterance was transcribed, and tagged for whether the parent, the child, and/or the experimenter, were the speaker and the addressee of said utterance. The target words were extracted from these utterances, and the target vowel boundaries were marked, allowing for selective formant analysis within these boundaries. Given the stimuli in Table 1, the placement of most vowel boundaries could be verified in the spectrogram by the voicing of the vowel as contrasted with the voiceless consonants surrounding it. In other cases the vowel is visually distinguishable from a voiced consonant by a burst, such as with a syllable-initial /b/, or by the changing formant pattern, such as for the syllable-final /n/ or syllable-initial /j/.

The recordings were annotated by undergraduate students with a basic background in linguistics, who are native speakers of Dutch and who received a task-specific training. This pre-processing of the data occurred independently of the current study; the steps described in §2.3-5 were all executed in the context of the study by Benders, StGeorge & Fletcher (2021). The steps described from this point onwards, in §2.6-2.8, were performed by the author of the current study.

2.6 TOKEN SELECTION

The vowels /i/, /u/, /a/, and /ɑ/ from the target words in Table 1 were included in analysis when the recording was clear, meaning free from sounds from the surroundings, and when there was no doubt about the utterance's addressee. Both of these properties were already marked in the corpus before the current analysis. Furthermore, tokens were excluded from analysis when neither the automatic formant trajectory extraction (as described in §2.7.1) nor the manual verification (as described in §2.7.3) could provide a conclusive analysis.

In total, 926 vowels were suitable for analysis, of which 145 instances of /i/, 289 of /u/, 180 of /a/, and 312 of /ɑ/, further specified in Tables 2 and 3. All speakers produced at least one token in each category and register.

Table 2: total number of tokens per vowel, broken down by gender and register.

	Infant-directed speech		Adult-directed speech		Total number of tokens per vowel
	Fathers	Mothers	Fathers	Mothers	
/i/	65	47	21	12	145
/u/	75	151	26	37	289
/a/	59	88	13	20	180
/ɑ/	138	118	23	33	312
Group total	337	404	83	102	926

Table 3: breakdown of number of tokens per vowel per participant group.

			Mean	Minimum	Median	Maximum
Infant-directed speech	Men	/i/	13	8	10	24
		/u/	15	3	15	25
		/a/	11.8	7	14	16
		/ɑ/	27.6	16	30	32
	Women	/i/	9.4	6	11	12
		/u/	30.2	8	35	61
		/a/	17.6	6	19	26
		/ɑ/	23.6	9	23	41
Adult-directed speech	Men	/i/	4.2	1	4	7
		/u/	5.2	1	7	10
		/a/	2.6	1	3	4
		/ɑ/	4.6	3	4	7
	Women	/i/	2.4	1	2	4
		/u/	7.4	4	6	16
		/a/	4	1	4	8
		/ɑ/	6.6	5	7	8

2.7 Acoustic measurements

For this study, the formant values for each token of /i/, /u/, /a/, and /ɑ/ were measured automatically in the speech analysis program Praat, then manually verified and corrected when necessary.

2.7.1 Formant measurement

The values for the first and second formant throughout each vowel were extracted automatically using the Burg algorithm as implemented in Praat (Boersma & Weenink 2022), and verified using a Praat-based interface scripted by Tuende Szalay and Titia Benders (Szalay 2022). The measurements were inspected visually in the spectrogram, and any artefacts found were corrected by adjusting the formant ceiling (as described in §2.7.2) or correcting the formant trajectory by hand (as described in §2.7.3). From the resulting formant contour, the median F1 and F2 in Bark (see §2.7.4) were taken from the central 40% of the vowel for each token. The median is selected rather than the mean, because it is robust to outliers, and the central 40% window is used to exclude variation caused by formant transitions at the vowel boundaries.

2.7.2 Formant ceilings

In order for a formant finding algorithm to find what it is meant to find, it needs to be given an appropriate frequency range in which to search for that value. Considering the differences between the sexes in the average size of the vocal tract and the consequent differences in the formant values in their speech, the upper limit of this window, the formant ceiling, differs between men and women. Additionally, due to the different formant values that characterize vowels, the optimal formant ceiling is slightly different for each vowel. The standard formant ceiling in Praat is 5500 Hz, which is reported by the developers to be the average formant ceiling for an adult female speaker. As the current study analyzes vowels one by one, extracted from longer interactions, and the gender of the speaker is known, a more targeted formant ceiling is appropriate. The current study uses the values specified in Table 4, as determined in Escudero et al. (2009). As these values were found from looking at Portuguese, there is no separate formant ceiling provided for /ɑ/, but as /ɑ/ is close to /a/, the formant ceiling for /a/ has been adopted here as well.

Table 4: optimal formant ceilings as determined by Escudero et al. (2009) in Hz, for each vowel, by gender.

	/i/	/u/	/a/	/ɑ/
Men	4910	4410	4340	4340
Women	6010	5200	5290	5290

2.7.3 MANUAL VERIFICATION OF FORMANT CONTOURS

In cases where the automatically extracted contour is unclear or irregular, adjustment of the formant ceiling can point the algorithm in the right direction. For example, in some cases where one or more formants are clearly being estimated too low (e.g. the F1 at around zero, or parts of the F2 equal to or below the F1 contour), the ceiling was raised incrementally until the contours of the measurements followed the formant contours in the spectrogram. Similarly, the ceiling was lowered if the estimated contour is placed higher than it should. The example in Figure 2 below shows an instance of /a/ as pronounced by a woman. The adopted standard formant ceiling of 4340 Hz from Table 4 for /a/ as pronounced by men does not produce three clear separate contours here, and in this specific example produced for illustration purposes, neither does the optimized ceiling of 5290 Hz for women. Increasing the formant ceiling in small steps, in this case to 6090 Hz, eventually results in clear formant contours for F1 to F3.

Another anomaly commonly found in formant contours in the analysis at hand was that just a very small subset of the points in the contour are off by a whole formant, higher or lower. In the second spectrogram in Figure 2, with the 5290 Hz formant ceiling, the first four sets of contour points from the left mark four formants, instead of three, and the first four points along the F3 contour are mistaken for the

F4. This also highlights the importance of considering the F3 in measurements, even when the F3 values are not used in further analysis. In the aforementioned interface for Praat, the formant contours are marked with separate colors for F1 to F3, which makes them more clearly distinguishable than in the example below. Small deviations are then easily corrected by hand.

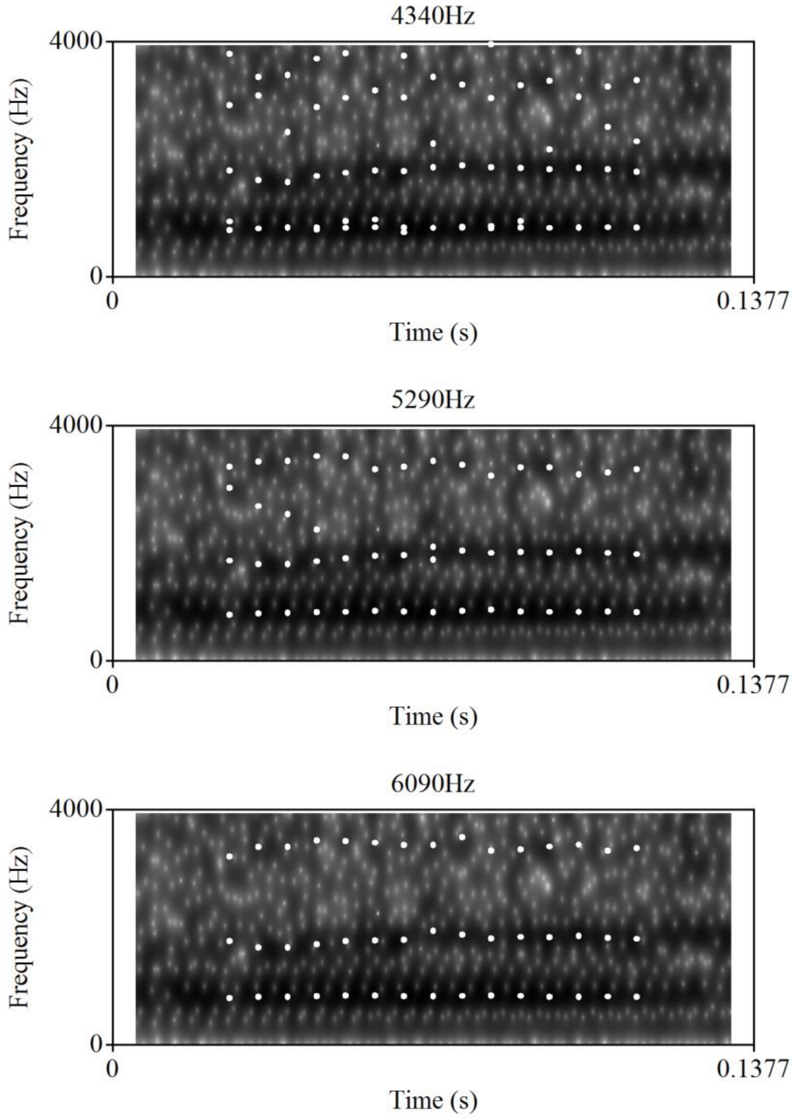


Figure 2: three analyses of the same /a/ as pronounced by an adult female, with the formant ceilings set as 4340 Hz, 5290 Hz, and 6090 Hz. All estimated formant contours are marked in white.

2.7.4 Conversion to Bark

The raw formant values in hertz were converted to Bark, prior to analysis. The psychoacoustic Bark scale (Zwicker 1961) was developed to more directly reflect the perceived frequency differences between sounds than the frequency of those sounds in hertz does. The formant values in Bark were calculated in Praat, which uses the equation as given in (1) (Boersma & Weenink 2022).

(1) $$Bark(x) = 7\log\left(\frac{Hz(x)}{650} + \sqrt{1 + \frac{Hz(x)}{650}^2}\right)$$

The current study uses Bark rather than hertz to be able to accurately draw comparisons between men and women by normalizing for the baseline differences in their formants.

2.8 Analysis

2.8.1 Vowel space size calculation

The size and shape of the vowel space used by the parents in infant-directed speech and adult-directed speech follow from the F1 and F2 measurements in the target vowels /i/, /a/, /ɑ/, and /u/. The vowel space size is computed for each group as the surface area of a simple polygon with the average F1 and F2 values in Bark for each of the four vowels as the *x*- and *y*-coordinates of its vertices. The surface area *A* of this vowel space polygon follows from the formula given in (2) (Habel et al. 2022).

(2) $$A = \left| \frac{x_1(y_4 - y_2) + x_2(y_1 - y_3) + x_3(y_2 - y_4) + x_4(y_3 - y_1)}{2} \right|$$

2.8.2 Unit of measurement

Analogous to the product of other units of measurement resulting in a surface area of said unit ‘squared’, or *unit*^{2} (e.g. $2\,m \cdot 2\,m = 4\,m^2$), the unit used for the vowel space size for the current study is *Bark*^{2}.

2.8.3 Software

All analyses were performed using R Statistical Software (version 4.2.0, R Core Team 2022) and RStudio (version 2022.02.3+492, RStudio Team 2022), with the *geometry* package (version 0.4.6, Habel et al. 2022) for vowel space size calculations, and the *magick* package (version 2.7.3, Ooms 2021) for imaging.

3. Results

The mean size and shape of the vowel space areas as produced by fathers and mothers in infant-directed speech and adult-directed speech are visualized below in Figure 3.

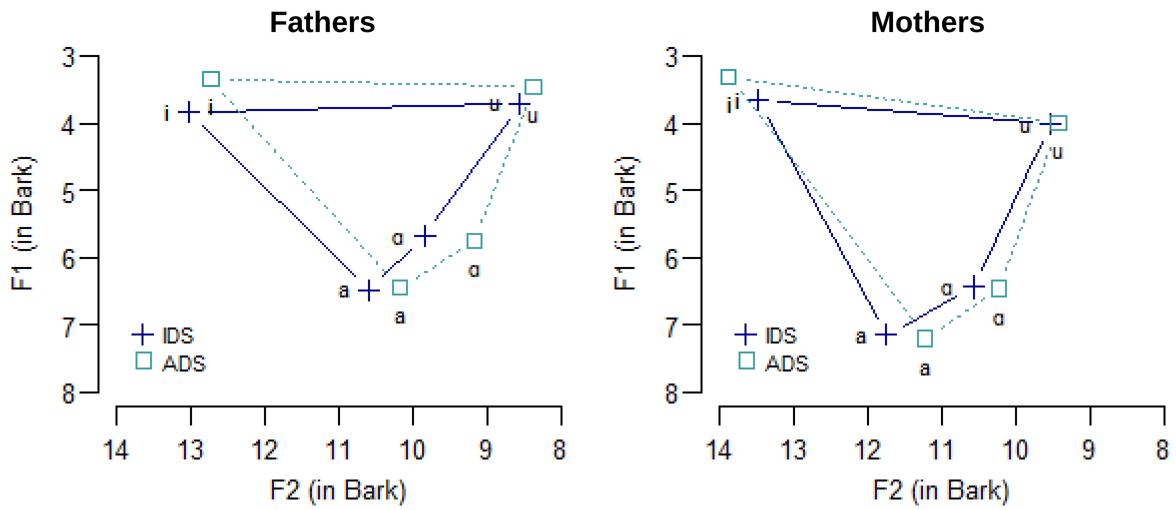


Figure 3: the vowel space area as used by fathers (left) and mothers (right) in infant-directed speech (marked with cross and solid line in dark blue) and in adult-directed speech (marked with square and dotted line in light blue).

The mean vowel space sizes are given in Figure 4 and Table 5. On average, the vowel spaces for both fathers and mothers are smaller in infant-directed speech, and the vowel spaces of fathers are smaller overall than those of mothers. While these means for each group suggest a main effect of both register and gender, a mixed ANOVA shows neither effect to be significant (register: $F(1,14) = 0.094$, $p = 0.763$; gender: $F(1,14) = 0.522$, $p = 0.482$). The means do not point towards an interaction effect, and the mixed ANOVA does not detect one either ($F(1,14) = 0.001$, $p = 0.972$).

The boxplots in Figure 5 show that the variance in the vowel space sizes, especially in infant-directed speech, is very large.

Table 5: size of the vowel space area in Bark², by gender and register.

			Mean	Minimum	Median	Maximum
Vowel space area	Fathers	Infant-directed speech	6.54	2.96	6.69	9.38
		Adult-directed speech	7.53	4.96	7.32	10.73
	Mothers	Infant-directed speech	7.59	4.65	6.46	13.36
		Adult-directed speech	8.47	7.68	8.44	9.82

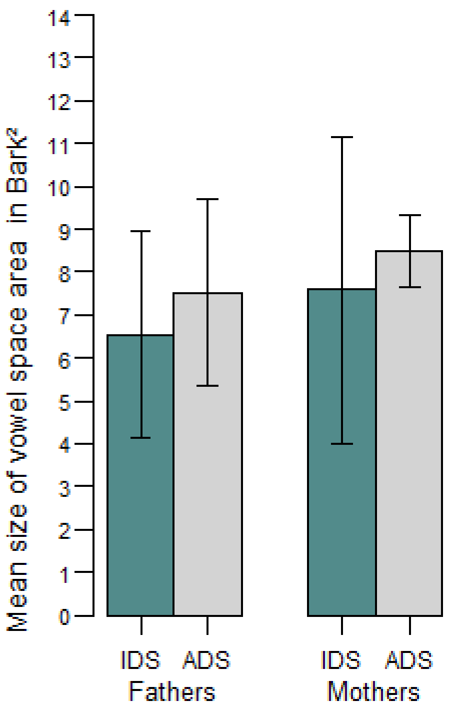


Figure 4: mean size of the vowel space area, by gender and by register (infant-directed speech in dark teal, adult-directed speech in light grey), error bar indicating standard deviation.

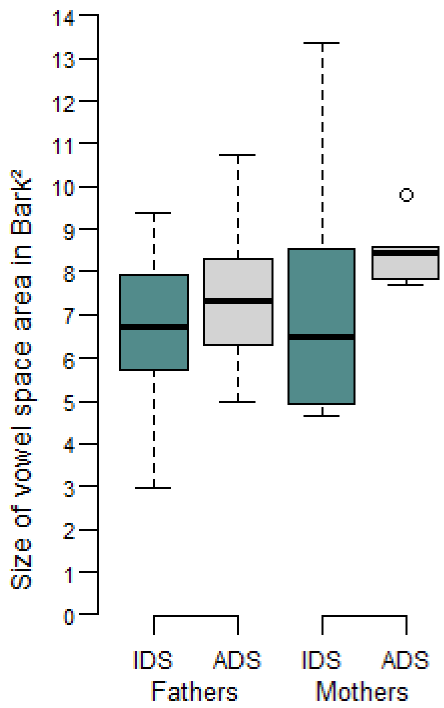
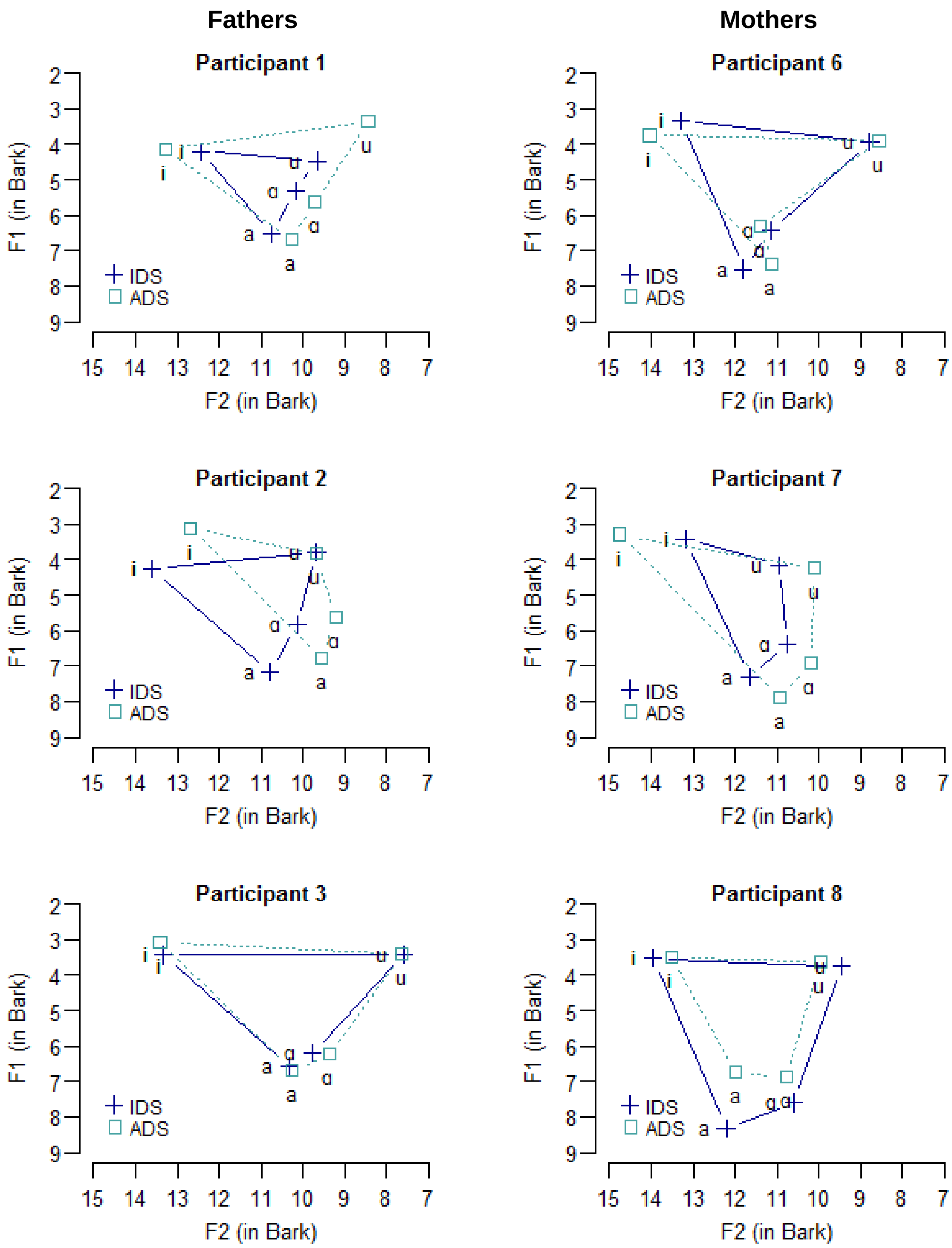


Figure 5: boxplots of the size of the vowel space area, by gender and register (infant-directed speech in dark teal, adult-directed speech in light grey).

The variance in shape and size of the vowel space for both groups is further illustrated by the individual vowel spaces for each participant given in Figure 6.



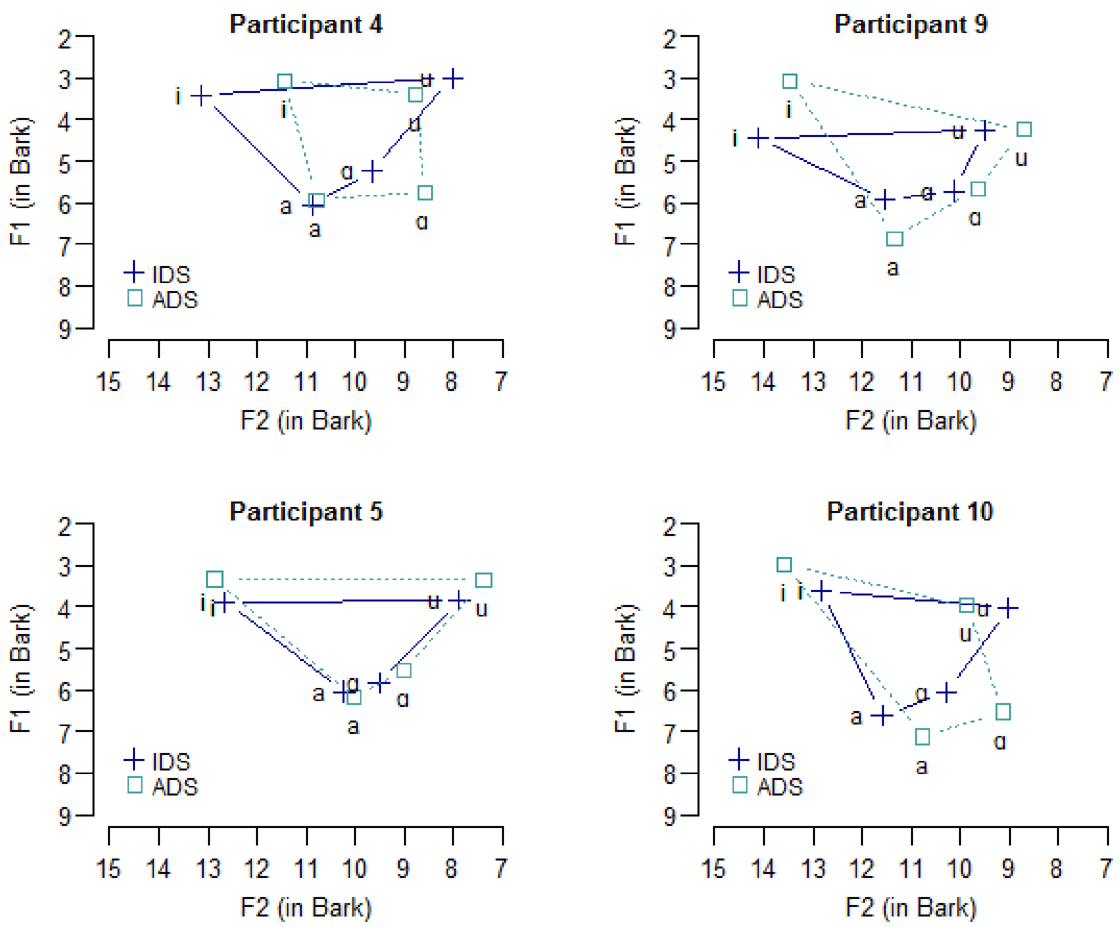


Figure 6: individually plotted vowel space areas of each participant in infant-directed speech (marked with cross and solid line in dark blue) and in adult-directed speech (marked with square and dotted line in light blue). Participants on the left (1 to 5) are fathers, participants on the right (6 to 10) are mothers.

4. Discussion

4.1 Results

4.1.1 Implications

The current study examined the size of Dutch fathers' vowel space in infant-directed speech, as compared to the size of their vowel space in adult-directed speech and compared to the vowel space sizes of mothers. The results show no evidence of an effect of register; fathers do not produce a smaller or larger vowel space with infants than with adults, and no effect of gender; fathers do not produce smaller or larger vowel spaces than mothers. No interaction effects were found. As no evidence for either hyper- or hypo-articulation was found, the results of the current study do not provide any basis for speculation on the general nature of either of these phenomena in infant-directed speech, or on the role of paternal speech as contrasted with the speech of mothers.

4.1.2 In light of previous research

It must be noted that between the many different methods and units of measurement and a general sparsity of information, there are currently no reference values available for the average vowel space size of speakers of Dutch. It is possible to compare the groups within the current study, with the parents' adult-directed speech serving as their own baseline, but aside from averages found by Benders (2013), the average vowel space sizes found in the current study cannot be compared to the typical vowel space for men and women.

In viewing the average values for the vowel space areas, it is remarkable that they differ substantially from the values found for Dutch mothers in Benders (2013), where means of 12.4 $Bark^2$ for infant-directed speech and 15.6 $Bark^2$ for adult-directed speech were reported. Proportionally, those are nearly double the values found for mothers in the current study, and the variance in the current study does not suggest the averages are attenuated because of a single downwards outlier. The method for measuring the formants was near-identical for both these studies, with the exception of the manual formant trajectory verification step added to the current study.

In order for this difference in method to be able to account for the large differences in average vowel space, some formant trajectories would have to have repeatedly been corrected in the same direction, leading to an overall smaller vowel space. For example, if the F1 for /a/ was reliably too high in the automatic formant estimation and consistently corrected in the manual verification step, that would have led to a smaller overall vowel space for all participants. While §2.7.2 does describe some artefacts that may occur in automatic formant measurements, no systematic anomalies of this type were detected at this time.

The main remaining observable difference between the studies is that the current study does have a much smaller set of participants. Where Benders analyzes the vowels of 18 mothers, the current study only considers 5. It is possible that this small number of participants and their large individual variance is skewing the data. After all, Benders' finding of a significantly smaller vowel space in infant-directed speech for Dutch mothers is not reproduced in the current study either.

4.2 LIMITATIONS OF THE CURRENT STUDY

4.2.1 PARTICIPANT SELECTION

The selection of participants for this study was limited by the composition of the available dataset. Because no background information was collected, a variety of factors that may be of influence on a parent's individual speech style in IDS cannot be controlled for in the current analysis. Factors could include (but are not limited to) parenting philosophy, socio-economic status, age, and education.

On parenting philosophy, for example, Shapiro, Hippe & Ramírez (2021) report that their male participants frequently reported being discouraged from ‘baby talk’ by their families, and they note that further research should include data collection on beliefs and attitudes. A review of studies on ‘motherese’ also concludes that the socio-economic status of mothers, as well as their general knowledge about child development, impact infant-directed speech (Saint-Georges et al. 2013). Regarding the method of participant recruitment for the original experiment, it is conceivable that the sample is biased for either one of these factors as compared to the general population. Considering that the participants volunteered, they are in a position to donate their time, which suggests they are not struggling financially, and they were also both interested in and aware of research activities in their area, which suggests a higher level of education. In Dutch, vowel quality patterns have been linked to age and education (Jacobi, Pols & Stroop 2006).

An unbalanced sample of participants does not by definition discount the results of a study, but if the sample is unbalanced, it is important to know about it and to be able to describe it, even if there are no conclusions or correlations related to it. And if there are, more background information on participants can be used to control for interactions and confounds.

4.2.2 Number of participants

In any scenario where including the parents' background in analysis is an option, a much larger number of participants would also be necessary. With the current sample size of $n = 10$, even if we had more information, finding connections between any of the factors mentioned and the use of vowel space in IDS would be anecdotal at best. Ideally, the sample size would be larger regardless of analysis of the parents' background, but as the current study is limited to the regular scope of a bachelor's thesis project, the samples used in the current study are smaller than preferred, and might not be representative of the group they are drawn from.

4.2.3 Age of the infant

The age of the infants addressed in infant-directed speech has been reported to be relevant to the parents' vowel space size in studies on American English and Hungarian (Cristia & Seidl 2014; Gergely et al. 2017) and irrelevant in studies on Australian English and Dutch (Kalashnikova & Burnham 2018; Benders 2013), so to err on the side of caution, the age window in the current study was set to be very narrow. The available dataset contains recordings of parents of infants in the 8-to-15-month range, and this study only selected parents from the 15-month-old group. While this does control for variation in vowel space use due to the age of the infant, it does not allow for a broad generalization about speech directed at infants at all ages that fall under the classification 'infant'. A more complete perspective may be gained by analyzing the vowel space use of parents in several infant age cohorts, as well as across languages. In cases where the vowel space size does indeed change with the age of the child, its increase or decrease may provide insight into the perceptual reasons for using an amended vowel space, and the relation between the vowel space and the different stages in children's language development.

4.2.4 Elicitation of speech

The main concerns in choosing a method for elicitation of speech are representation and comprehensiveness: is the sample representative of the type of speech we are analyzing, and is does it contain all the elements necessary for a complete analysis? The physical location and the social setting in which the participants are recorded is of influence on their speech. Various activities call for separate speech patterns, and differences include (but are not limited to) types of utterances, intonation patterns, vocabulary, and affect or dramatization.

In the current study, taking a pre-selected set of items out of a bag supplied each interaction with certain target words resulting in a number of vowel tokens suitable for analysis. Table 2 however does show that the sample is not necessarily balanced – there were far more tokens for /ɑ/ and /u/ than the other two vowels. A more spontaneous interaction may possibly better reflect natural speech outside of a laboratory setting, but will also produce varied samples that may not contain the target vowels in a selection of words suitable for analysis. In Dutch, for example, sentence accent, word stress, and word class have been shown to affect the pronunciation of vowels, leading to different formant patterns and vowel duration (van Bergem 1993). In order to be able to properly compare the vowels of different speakers, it is important that they occur in the same, or highly similar words and sentences, to rule out any variation caused by such differences in the composition of the speech sample.

The activity used to elicit the dataset at hand is not entirely unlike what a play session at home might look like[1], so rather than analyzing non-elicited spontaneous speech, an analysis more comprehensive than the current study might include speech from multiple types of settings that occur naturally between the parent and the child.

4.2.5 Research methods and standards

In largely following the method of Benders (2013), the current study did not consider other units of measurement used in measuring vowel space size. While the choices for the current study are elaborately motivated in §2.4.7 and §2.8.2, there are other methods of normalization, and it would be interesting to see them contrasted. Considering the discussion in §4.1.2, future research may benefit from a closer look at differences that originate in differences in method of measurement and unit of measurement. While the current study does not intend to draw conclusions from a comparison of results with other studies, it would be especially important to control for variation caused by different methods in any future cases where vowel space sizes are compared between studies.

A common theme between the points of discussion in this section and in those in §4.1.2–4.2.1 and §4.2.4 is the wide variety of available methodological choices and consequent compatibility issues when building on previous findings. There is no direct reason to claim the entire area of infant-directed speech vowel space research would

[1] Anecdotally evidenced by the author by many years of working with small children and unpacking many a bag, basket, or cupboard filled with toys or household items.

benefit from more standardization in general, because in asking *can we generalize*, the immediately following question is *should we want to*? For many research questions, there is no obvious reason to compare the values found to those found in other studies. For example, whether the register determines the size of the vowel space can perfectly be examined within-speaker, with no need for outside reference values. This example concerns a well-delineated question which can be answered on its own, and the methodology may be perfectly suited towards that. However, the answer to this question does also carry the potential to contribute more towards the broader field of infant-directed speech research, if it could be considered in its context by others. In order to be able to do that, the methodology and any values found need to be reported on extremely transparently. In constructing the theoretical framework of this study, a common obstacle was studies which only reported on the vowel space visually. This statement begs for an example, but none of these studies are referenced here because they were rejected in the theoretical research phase of this study for being unhelpful. In summary, rather than making an argument for standardization of methods, consider this a call for transparency in reporting.

As a general closing remark, it is prudent that future research include male parents and caregivers in research into child-directed speech and adult contributions to child language development. A wide range of features of infant-directed speech which are correlated with child language outcomes which have been studied in women, but not in men, remain open to investigation.

5. Conclusion

The current study investigated the vowel space size of Dutch fathers in infant-directed speech. No evidence was found that fathers amend their vowel space in infant-directed speech, as compared to their articulation in adult-directed speech, and as compared to mothers. Further studies will benefit most from a more substantial sample size. Other recommendations for future research include consideration of data on the background of participants, and increased transparency on method of vowel space measurement.

Acknowledgements

First and foremost I must thank Titia Benders, whose previously collected dataset made it possible for me to investigate this topic, and whose feedback and support have been valuable beyond measure. I would also like to thank Suki Yiu, for her enthusiasm in the weeks leading up to the thesis trajectory, as well as for being my second reader. My peer group mates Issy, Ceyda, Long, and Lieke: it's been an honor working alongside you guys. And finally, I owe many thanks to the parents in the recordings and the annotators, whom I have never met but without whom this thesis would not be before you today.

References

Benders, Titia. 2013. Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behavior and Development* 36(4). 847–862. https://doi.org/10.1016/j.infbeh.2013.09.001.

Benders, Titia, Jennifer StGeorge & Richard Fletcher. 2021. Infant-directed Speech by Dutch Fathers: Increased Pitch Variability within and across Utterances. *Language Learning and Development* 17(3). 292–325. https://doi.org/10.1080/15475441.2021.1876698.

Boersma, Paul & David Weenink. 2022. *Praat: doing phonetics by computer*. http://www.praat.org/.

Byers-Heinlein, Krista, Angeline Sin Mei Tsui, Christina Bergmann, Alexis K. Black, Anna Brown, Maria Julia Carbajal, Samantha Durrant, et al. 2021. A Multilab Study of Bilingual Infants: Exploring the Preference for Infant-Directed Speech. *Advances in Methods and Practices in Psychological Science*. SAGE Publications Inc 4(1). 2515245920974622. https://doi.org/10.1177/2515245920974622.

Cox, Christopher, Christina Bergmann, Emma Fowler, Tamar Keren-Portnoy, Andreas Roepstorff, Gregory Bryant & Riccardo Fusaroli. 2022. *A Systematic Review and Bayesian Meta-Analysis of the Acoustic Features of Infant-Directed Speech*. Preprint. In Review. https://doi.org/10.21203/rs.3.rs-1396005/v1.

Cristia, Alejandrina. 2013. Input to Language: The Phonetics and Perception of Infant-Directed Speech. *Language and Linguistics Compass* 7(3). 157–170. https://doi.org/10.1111/lnc3.12015.

Cristia, Alejandrina & Amanda Seidl. 2014. The hyperarticulation hypothesis of infant-directed speech*. *Journal of Child Language*. Cambridge University Press 41(4). 913–934. https://doi.org/10.1017/S0305000912000669.

Englund, Kjellrun & Dawn Behne. 2006. Changes in infant directed speech in the first six months. *Infant and Child Development* 15(2). 139–160. https://doi.org/10.1002/icd.445.

Englund, Kjellrun T. 2018. Hypoarticulation in infant-directed speech. *Applied Psycholinguistics* 39(1). 67–87. https://doi.org/10.1017/S0142716417000480.

Escudero, Paola, Paul Boersma, Andréia Schurt Rauber & Ricardo A. H. Bion. 2009. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *The Journal of the Acoustical Society of America* 126(3). 1379–1393. https://doi.org/10.1121/1.3180321.

Fernald, Anne. 1985. Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development* 8(2). 181–195. https://doi.org/10.1016/S0163-6383(85)80005-9.

Gergely, Anna, Tamás Faragó, Ágoston Galambos & József Topál. 2017. Differential effects of speech situations on mothers' and fathers' infant-directed and dog-directed speech: An acoustic analysis. *Scientific Reports* 7. https://doi.org/10.1038/s41598-017-13883-2.

Gussenhoven, Carlos. 1992. Dutch. *Journal of the International Phonetic Association* 22. 45–47. https://doi.org/10.1017/S002510030000459X.

Habel, Kai, Raoul Grasman, Robert B. Gramacy, Pavlo Mozharovskyi & David C. Sterratt. 2022. *geometry: Mesh Generation and Surface Tessellation*. https://CRAN.R-project.org/package=geometry. (12 June, 2022).

Jacobi, I. 2009. *On variation and change in diphthongs and long vowels of spoken Dutch*. https://dare.uva.nl/search?identifier=39bfa671-db0f-4725-959e-83b8bec183ad. (27 June, 2022).

Jacobi, I., L. C. W. Pols & J. Stroop. 2006. Measuring and Comparing Vowel Qualities in a Dutch Spontaneous Speech Corpus. *Interspeech* 7. https://dare.uva.nl/search?identifier=327061a6-f6df-49bf-9684-f569bc1e4e0e. (27 June, 2022).

Jong, Nivja H. de & Ton Wempe. 2009. Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods* 41(2). 385–390. https://doi.org/10.3758/BRM.41.2.385.

Kalashnikova, Marina & Denis Burnham. 2018. Infant-directed speech from seven to nineteen months has similar acoustic properties but different functions. *Journal of Child Language* 45(5). 1035–1053. https://doi.org/10.1017/S0305000917000629.

Kalashnikova, Marina & Manuel Carreiras. 2021. Input quality and speech perception development in bilingual infants' first year of life. *Child Development* 93. https://doi.org/10.1111/cdev.13686.

Kempe, Vera, David A. Puts & Rodrigo A. Cárdenas. 2013. Masculine Men Articulate Less Clearly. *Human Nature* 24(4). 461–475. https://doi.org/10.1007/s12110-013-9183-y.

Kuhl, Patricia K., Jean E. Andruski, Inna A. Chistovich, Ludmilla A. Chistovich, Elena V. Kozhevnikova, Viktoria L. Ryskina, Elvira I. Stolyarova, Ulla Sundberg & Francisco Lacerda. 1997. Cross-Language Analysis of Phonetic Units in Language Addressed to Infants. *Science*. American Association for the Advancement of Science 277(5326). 684–686. https://doi.org/10.1126/science.277.5326.684.

Lam, Christa & Christine Kitamura. 2010. Maternal Interactions With a Hearing and Hearing-Impaired Twin: Similarities and Differences in Speech Input, Interaction Quality, and Word Production. *AustraliaJournal of Speech, Language, and Hearing Research* 53. 543–555.

Liu, Huei-Mei, Patricia K. Kuhl & Feng-Ming Tsao. 2003. An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science* 6(3). F1–F10. https://doi.org/10.1111/1467-7687.00275.

Marklund, Ulrika, Ellen Marklund & Lisa Gustavsson. 2021. Relationship Between Parent Vowel Hyperarticulation in Infant-Directed Speech and Infant Phonetic Complexity on the Level of Conversational Turns. *Frontiers in Psychology* 12. 688242. https://doi.org/10.3389/fpsyg.2021.688242.

Ooms, Jeroen. 2021. *magick: Advanced Graphics and Image-Processing in R*. https://CRAN.R-project.org/package=magick. (12 June, 2022).

Pancsofar, Nadya & Lynne Vernon-Feagans. 2006. Mother and father language input to young children: Contributions to later language development. *Journal of Applied Developmental Psychology* 27(6). 571–587. https://doi.org/10.1016/j.appdev.2006.08.003.

R Core Team. 2022. *R: A language and environment for statistical computing*. Vienna, Austria. https://www.R-project.org/.

RStudio Team. 2022. *RStudio: Integrated Development Environment for R.* Boston, MA: RStudio, PBC. http://www.rstudio.com/.

Saint-Georges, Catherine, Mohamed Chetouani, Raquel Cassel, Fabio Apicella, Ammar Mahdhaoui, Filippo Muratori, Marie-Christine Laznik & David Cohen. 2013. Motherese in interaction: at the cross-road of emotion and cognition? (A systematic review). *PloS One* 8(10). e78103. https://doi.org/10.1371/journal.pone.0078103.

Shapiro, Naomi Tachikawa, Daniel S. Hippe & Naja Ferjan Ramírez. 2021. How Chatty Are Daddies? An Exploratory Study of Infants' Language Environments. *Journal of Speech, Language, and Hearing Research* 64(8). 3242–3252. https://doi.org/10.1044/2021_JSLHR-20-00727.

Szalay, Tuende. 2022. Prelateral vowel merger in Australian English: /æɔ-æ/ [Manuscript submitted for conference]. *The University of Sydney*.

The ManyBabies Consortium. 2020. Quantifying Sources of Variability in Infancy Research Using the Infant-Directed-Speech Preference. *Advances in Methods and Practices in Psychological Science*. SAGE Publications Inc 3(1). 24–52. https://doi.org/10.1177/2515245919900809.

Uther, M., M. A. Knoll & D. Burnham. 2007. Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication* 49(1). 2–7. https://doi.org/10.1016/j.specom.2006.10.003.

Weirich, Melanie & Adrian P. Simpson. 2018. Gender identity is indexed and perceived in speech. *PLOS ONE*. Public Library of Science 13(12). e0209226. https://doi.org/10.1371/journal.pone.0209226.

Xu, Nan, Denis Burnham, Christine Kitamura & Ute Vollmer-Conna. 2013. Vowel Hyperarticulation in Parrot-, Dog- and Infant-Directed Speech. *Anthrozoös* 26(3). 373–380. https://doi.org/10.2752/175303713X13697429463592.

Zwicker, E. 1961. Subdivision of the Audible Frequency Range into Critical Bands (Frequenzgruppen). *The Journal of the Acoustical Society of America* 33(2). 248. https://doi.org/10.1121/1.1908630.