Effects of gender, parental role and speech material on infant-directed speech in Swedish

Melanie Weirich[1], Adrian P. Simpson[1], Christine Ericsdotter Nordgren[2] & Jasmine Öjbro[2]
[1]Germanistische Sprachwissenschaft, FSU Jena, Germany [2]Språkstudion vid Institutionen för språkdidaktik, Stockholms universitet, Sweden

ABSTRACT

While phonetic features of infant-directed speech (IDS) have been described cross-linguistically, especially in mothers, the reasons and conditions under which they are found are less clear. In this study, phonetic cues of IDS are investigated in 21 Swedish mothers and fathers including mean fundamental frequency (f0), variation in f0, formant values and vowel space size. The speech material contains read and semi-spontaneous speech. Also, the involvement in child care is considered as a potential interacting factor.

Results show that differences between mothers and fathers in IDS are small and that differences in the involvement in child care are not reflected in the extent of the IDS features used. F0 differences are found in IDS in both semi-spontaneous speech and (although to a smaller extent) in read speech. In contrast, vowel space enhancement is only found in semi-spontaneous speech and differences between the genders appear regarding individual formants of certain vowels.

Keywords: infant-directed speech, gender, fundamental frequency, vowel space size, formants

1. INTRODUCTION

Infant-directed speech (IDS) or motherese has been found to be characterized cross-linguistically by various phonetic modifications, such as increased average fundamental frequency (f0), larger f0 excursions ([1] for French, Italian, German, Japanese, British English, and American English, [2] for British English) and enhanced vowel space (e.g. [3,4] for Russian, Swedish and English, [5] for Mandarin Chinese). In connection with these findings, the question has been discussed whether vowel space expansion is indeed an inherent characteristic of IDS with the purpose of enhancing acoustic contrasts between sounds, thus promoting language acquisition ([4,6]). [5] on Mandarin Chinese supports this assumption showing that a relationship exists between mothers' use of enhanced vowel spaces in IDS and their children's speech discrimination performance. Similarly, [7] found that vowel hyperarticulation enhances word recognition in 19-month-olds. While vowel hyperarticulation is also found in speech to foreigners ([6]) or to hearing impaired adults ([8]), it is not found in pet-directed speech ([9]). This indicates that hyperarticulation may be related to the talker's knowledge of the listener's linguistic competence.

Other studies have suggested that vowel space expansion is not a result of hyperarticulation per se but rather results from formant raising as a byproduct of a) trying to sound smaller and non-threatening acoustically indexed by higher formants in general ([10]), or b) increased smiling behavior in IDS ([11,12,13]). This claim is based on the lack of finding enhanced phonetic contrasts in IDS ([11,12,14,15,16]) and the importance of the infant's feedback. [17] tested the role of the infants' feedback by manipulating the mothers' knowledge of whether they believed their infants could hear them or not, and by changing the audibility of the speech signal available to the infant. While the mothers' knowledge did not affect the vowel articulation, the infant's feedback did: enhanced vowel spaces were only found in the audibility condition where infants gave feedback to the mothers' interaction. [13] tested the importance of feedback by using storybook read speech, where mothers focused their attention on the book rather than on their infants. And, indeed, they did not find an expansion of the vowel space but only a systematic shift of the first two formants in /i/ and /u/. The present study follows this line of research, investigating whether IDS shows enhanced phonetic contrasts to facilitate speech acquisition or whether the infant's feedback/communication with the adult is the triggering factor. Going a step further than [13], the relevance of the communicative situation is investigated not only by looking at read speech, but by comparing IDS in read and semi-spontaneous speech elicited from the same participants.

Furthermore, IDS has been investigated primarily in mothers, who traditionally are the main caregivers. Studies including fathers show differences between the genders regarding prosodic modifications (cf. [2]). A possible interacting factor might be the involvement in child care, as was suggested in [18], where a relationship between the involvement of a father and the amount of speech directed to the child

was found. Regarding vowel space size, we are not aware of any larger study including fathers. Thus, the present study comprises read and semi-spontaneous speech of both mothers and fathers; the acoustic parameters investigated are f0 modifications, formants and vowel space size. Additionally, the involvement in child care is examined as a potential influencing factor.

2. METHOD

The data presented here is part of a larger longitudinal project ([19,20]) investigating IDS and ADS in Swedish and German mothers and fathers, also examining the involvement in child care and the self-ascribed gender identity of a speaker on IDS and ADS over a time span of one year. The fathers in this material differ with respect to timing and extent of parental involvement, while the mothers constitute a much more homogeneous group in this regard. Speech recordings are made with the participating caregiver before the child is born and at three time points during the child's first year. In this paper, Swedish IDS (directed to the infant) and ADS (directed to the experimenter) of the second recording are presented. The age of the babies ranges between 4 months, 5 days and 6 months, 27 days. To account for individual differences between speakers and genders, for each speaker, the analyzed parameters of the first recording are taken as reference values (and set to 100%). Values at time 2 are calculated in relation to the reference value.

2.1 Participants, speech material and involvement in child care

The data comprises recordings at two time points of 10 mothers (mean age: 30.8) and 11 fathers (mean age: 31.5). Participants were recorded in their homes by the same female experimenter using a headset-microphone (Sennheiser ew 100 G3 – SK100) and a ZOOM – H6 Handy Recorder. The speech material consists of a) a semi-spontaneous picture describing task using 15 pictures showing, among other things, the carrier words, e.g., *ko* (cow), *bok* (book), *bord* (table), *väska* (bag), *katt* (cat), *fisk* (fish), *himmel* (sky), *kaka* (cake) and b) read speech using different short texts of the book "Alla vi barn i Bullerbyn" ([21]). The texts were slightly changed to contain the target vowels within recurring carrier words such as the names of the children and other common words (e.g. *Lasse, Bettan, Rolf, flicka* (girl), *fick* (got). Table 1 gives an overview of the number of tokens investigated separated by speech material and register.

In addition to the recordings, data was gathered relating to involvement in child care using a questionnaire. Participants were asked about (a) the amount of time spent with the child and the tasks done with the child (12 items) and (b) the amount of speech (singing, reading, talking) directed to the child (3 items). Items were rated on a scale from 1 (seldom/little/always the partner) to 7 (often/much/ always myself). Two average values were then calculated for each participant reflecting their general involvement in child care (a) and the amount of speech directed to the child (b).

Table 1. Number of tokens separated by speech material and speech register (ADS/IDS)

	Read speech (ADS/IDS)	Picture description (ADS/IDS)
i	209/122	246/106
ɛ	144/81	357/193
a	215/128	202/111
ɑ	-	142/76
ɔ	199/113	-
u	198/117	607/280

2.2 Acoustic analysis

Formants and fundamental frequency measures were analyzed in PRAAT ([22]). F1 and F2 were estimated at the vowel midpoint and mean values for each speaker/register and speech task were calculated (in Bark). Vowel space sizes were estimated using the polygon in F1xF2 space defined by the vowels /i ɛ a ɔ u/ for read speech and /i ɛ a ɑ u/ for the picture description. Also, mean fundamental frequency (f0) and variation in f0 (SD) were calculated (range: 50-600 Hz) for each speaker/register and speech task. Any confounding vocalization of the child was cut from the recording.

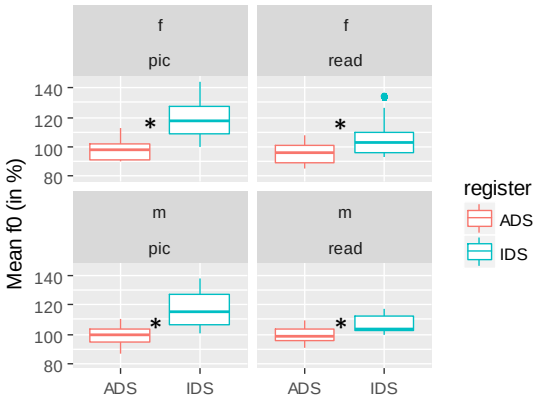
3. RESULTS

Linear mixed models were run using the lme4 package ([23]) in R ([24]) separately for read and spontaneous speech and each acoustic parameter. We entered speaker gender (m, f) and the speech register (ADS, IDS) as potential fixed factors and a random intercept for speaker. P-values were obtained using Likelihood ratio tests comparing the model with the factor/interaction in question with the model without this factor/interaction. Also, Pearson correlations were run between the parental involvement/amount of speech directed to the child and the degree of IDS (i.e. increase in vowel space or f0 in %).

3.1 Fundamental frequency

Figure 1 shows mean f0 as a function of register, separated by speech material and speaker gender. Remember that values are given in %, expressed in relation to each speaker's reference measurements at the first time point to normalize between individuals and genders. Differences between the registers can be seen in the figure and these are reflected in the statistical analysis. For the picture description task, a significant main effect for register was found ($\chi^2(1)$= 20.5, $p < .001$) reflecting increased mean f0 in IDS compared to ADS regardless of speaker gender. For read speech, also, a significant effect of register was found ($\chi^2(1)$= 21.4, $p < .001$), even though the difference between IDS and ADS is smaller here than in the semi-spontaneous speech (Estimate = 9.4 percent points for read speech vs. 22.1 for picture description). Note that results do not change when the Bark scale is used.

Figure 1. Mean fundamental frequency (f0, in % of reference recording at time 1) separated by speech material, speaker gender and speech register. Significant differences marked by an asterisk.

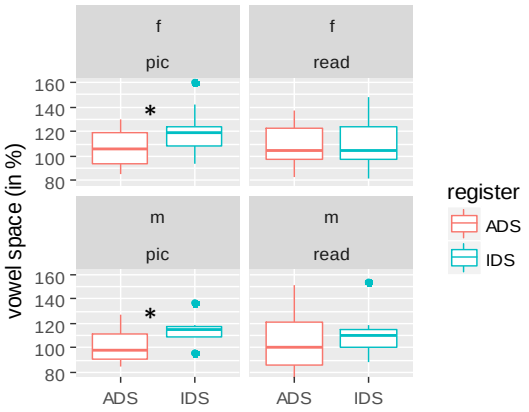


Also regarding variation in f0 (SD) a significant effect of register was found for both the picture description task ($\chi^2(1) = 7.9$, $p < .01$) and read speech ($\chi^2(1) = 4.4$, $p < .05$). Here, too, the increase in f0 variation in IDS was stronger in the picture description task with 30.4 percent points compared to only 15.1 in the read speech.

3.2 Vowel space size

Figure 2 shows the vowel space size (again in % of time 1) separated by speech material, gender and register. Differences between IDS and ADS can be seen for the picture description task and a significant main effect of register was found ($\chi^2(1)$= 15.2, $p <$ .001). There was no difference between fathers and mothers. Thus, as apparent from the figure, both mothers and fathers show significantly larger vowel spaces in IDS than in ADS. Even though a tendency for increased vowel spaces in fathers can be seen also for read speech, this effect was not significant in the model, probably due to the high variability within ADS.

Figure 2. Variation in vowel space size (in % of reference recording at time 1) separated by speech material, speaker gender and speech register. Significant differences marked by an asterisk.



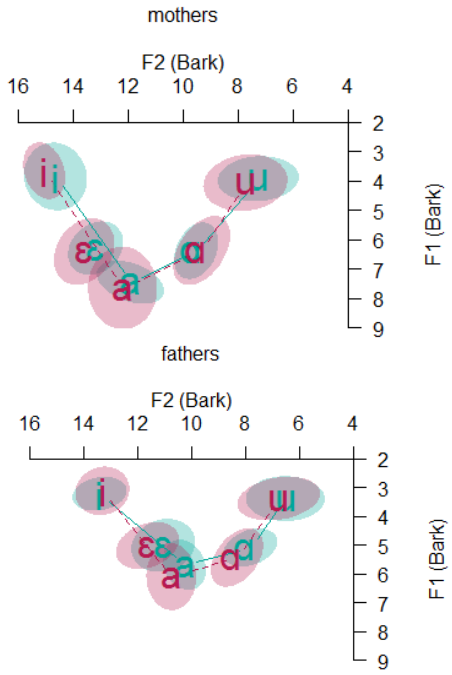
To further investigate the differences between the individual vowel categories Figure 3 shows the average vowel spaces for IDS (red, dashed line) and ADS (blue, solid line) of the picture description task separately for fathers and mothers. The different sizes of the ellipses reflect the difference in intra- and inter-speaker variation between the vowel categories. It is apparent from the figure that there is no general increase in formant values across all vowels and both F1 and F2, but rather individual changes in particular formants and vowels. In detail, /a/ and /ɛ/ (and in fathers also /ɑ/) show higher F1 and F2 values in IDS, thereby shifting the vowel space towards a more fronted and open articulation. While /i/ in fathers stays stable, it shows higher F2 but lower F1 values in mothers, resulting also in a more fronted articulation but also an increased distance to /a/ in IDS compared to ADS. /u/ behaves somewhat differently between the genders: while in fathers, only a change in F2 can be seen, in mothers both F1 and F2 reveal higher values.

The statistical analysis with F2 as dependent variable showed a significant interaction of vowel category and register in mothers ($\chi(5)$=19.6, $p < .01$), with increased F2 values in IDS in /i/ ($p < .05$), /u/ ($p <$ .01) and /ɛ/ ($p < .01$). For fathers on the other hand, a significant interaction of vowel and register was found for F1 ($\chi(5) = 27.45$, $p < .001$) with increased values in IDS for /a/ ($p < .001$) and /ɑ/ ($p<.01$). Thus, vowel and formant specific changes are responsible for the larger vowel space size in IDS in spontaneous speech. Results thereby do not corroborate the existence of a larger vowel space in IDS being due to an increased smiling behavior or due to a lengthening of the vocal tract and the effort to

sound smaller and non-threatening [10]). In addition, the changes differ between the genders.

For read speech, a significant main effect of register was found for F2 in mothers (χ (1) = 4.02, $p < .05$) and no interaction with vowel category, mirroring a higher F2 value in IDS regardless of vowel. Also, in fathers no interaction between formant value and vowel category was found. However, the main effect of register on F1 only marginally failed to show significance ($p = .057$). Overall, the effect of register on vowel formants seems to be more general in the read speech task than in semi-spontaneous speech pointing to different underlying mechanisms, i.e. in read speech we find a shift in vowel space (more fronted in females, more open in males), whereas in spontaneous speech we find an increase in vowel space size.

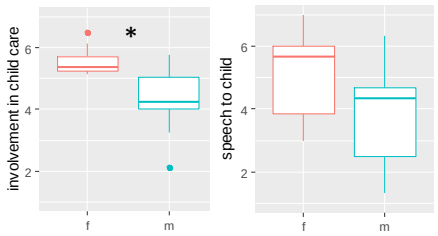
Figure 3. Average vowel spaces (in Bark) in ADS (solid line, blue) and IDS (dashed line, red) of the picture description task separated by speaker gender.



3.3 Involvement in child care

Figure 4 shows the distribution of the involvement in child care (left plot) and the amount of speech directed to the child (right plot) separated by speaker gender. As expected, mothers reveal higher values than fathers with respect to the involvement in child care ($p < .001$). The difference between the genders in the amount of speech directed to the child was not significant ($p=.10$). Within the fathers, a high degree of variation appears regarding both scales, with some fathers showing very small values and others showing values similar to the mothers. Within the mothers, variation is also high regarding the amount of speech directed to the child. This variation in child care involvement/speech to child, however, was not found to be related to any of the IDS features. No significant correlation between the involvement in child care or the amount of speech directed to the child and the degree of IDS with respect to any of the acoustic parameters was found within mothers or fathers.

Figure 4. Variation in involvement in child care (left) and speech to child (right) separated by speaker sex. Significant differences marked by an asterisk.



4. DISCUSSION

The Swedish mothers and fathers of our study behaved similarly regarding the acoustic-phonetic features found in IDS but involvement in child care was not an explaining factor. A potential reason for this might be the ongoing decrease in gender differences in parenting. For example, in the 1990s gender differences found in vocabulary and conversational aspects of IDS have been claimed to be due to differing parental roles, with mothers supposed to "provide a feeling of security" vs. fathers supposed to "prompt the child to attain higher levels of success" [25]. This, however, seems to be outdated especially in a society like Sweden with a leading position in gender equality and compatibility of family and work in Europe since the 1970s.

A clear effect of speech register was found on mean f0 and variation in f0 for both speech tasks. The effect was however larger for the picture description than for the read speech. Larger vowel space sizes were found in IDS only for the picture description task. Thus, we did indeed find that IDS characteristics were stronger in semi-spontaneous than in read speech. This is in line with studies emphasizing the importance of a communicative situation and the infant's feedback [13]. In addition, the underlying formant changes differed between the speech tasks and, also between the genders. While fathers showed stronger effects for F1, mothers showed larger differences in F2. These effects were vowel specific in the semi-spontaneous speech but not in the read speech, resulting in larger vowel spaces only in the first one. Thus, while keeping in mind that speaker sample is small, neither a general increase in formants reflecting the adult's effort to sound non-threatening [10], nor an increase in F2 due to increased smiling behavior [11] can explain the vowel space enhancement found in spontaneous IDS.

5. REFERENCES

[1] 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, 477–501.

[2] Shute B & Wheldall K (1999) Fundamental frequency and temporal modifications in the speech of british fathers to their children. Educational Psychology 19(2), 221-233.

[3] Sundberg U (1998) Mother tongue - Phonetic Aspects of Infant-Directed Speech. PhD dissertation. Stockholm University: Department of Linguistics.

[4] Kuhl PK, Andruski JE, Chistovich IA, Chistovich LA, Kozhevnikova EV, Ryskin VL, Stolyarova, EI, Sundberg U & Lacerda F (1997) Cross language analysis of phonetic units in language addressed to infants. Science, 277(5326), 684–686.

[5] Liu H-M, Kuhl PK & Tsao F-M (2003) An association between mothers' speech clarity and infants' speech discrimination skills. Developmental Science, 6 (3), F1–F10.

[6] Uther M, Knoll MA & Burnham D (2007) Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. Speech Communication, 49 (1), 2–7.

[7] Song JY, 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, 389–400.

[8] Ferguson SH & Kewley-Port D (2002) Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. Journal of the Acoustical Society of America, 112 (1), 259–271.

[9] Burnham D, Kitamura C & Vollmer-Conna U (2002) What's new pussycat: on talking to animals and babies. Science, 296, 1435.

[10] Kalashnikova M., Carignan C., Burnham D. (2017). The origins of babytalk: smiling, teaching or social convergence? Royal Society open science, 4, 170306

[11] Benders T (2013) Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. Infant Behavior & Development 36, 847-862.

[12] Cristia A & Seidl A (2013) The hyperarticulation hypothesis of infant directed speech. Journal of Child Lan-guage 40 (2), 1-22.

[13] Burnham EB, Wieland EA, Kondaurova MV, McAuley JD, Bergeson TR & Dilley LC (2015) Phonetic Modification of Vowel Space in Storybook Speech to Infants up to 2 Years of Age. Journal of Speech Language and Hearing Research. 58, 241-253.

[14] Englund K & Behne D (2006) Changes in infant directed speech in the first six months. Infant and Child Development 15, 139–160.

[15] Kondaurova MV, Bergeson TR & Dilley LC (2012) Effects of deafness on acoustic characteristics of American English tense/lax vowels in maternal speech to infants. Journal of the Acoustical Society of America, 132(2), 1039-1049.

[16] Martin A, Schatz T, Versteegh M, Miyazawa K, Mazuka R, Dupoux E. & Cristia A. (2014) Mothers Speak Less Clearly to Infants Than to Adults: A Comprehensive Test of the Hyperarticulation Hypothesis. Psychological Science, 1–7.

[17] Lam C & Kitamura C (2012) Mommy, speak clearly: induced hearing loss shapes vowel hyperarticulation. Developmental Science 15(2), 212–221

[18] Sheehan E A (2004) Influence of paternal involvement on fathers' infant-directed speech and infants' brain activity to male and female speech. PhD dissertation. Emory University.

[19] Weirich, M. & Simpson, A. (2016) Changes in IDS and ADS during parental leave - project sketch and first results of pilot studies. Proceedings of the 12th meeting on "Phonetik und Phonologie im deutschsprachigen Raum" (P&P12), Munich, 224-227.

[20] Weirich, M. & Simpson, A. (2017) Acoustic correlates of parental role and gender identity in the speech of expecting parents, Interspeech, Stockholm, 924-928.

[21] Lindgren, A. (1947) Alla vi barn i Bullerbyn. Rabén & Sjögren, Stockholm.

[22] Boersma P & Weenink D (2012) Praat: doing phonetics by computer [Computer program]. Retrieved from http://www.praat.org.

[23] Bates D, Maechler M, Bolker B & Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4, Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.

[24] R Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[25] Le Chanu, M. & Marcos, H. (1994) Father-Child and Mother-Child Speech: A perspective on parental roles. European Journal of Psychology of Education, 9(1), 3-13.