INTERSPEECH 2017
August 20–24, 2017, Stockholm, Sweden

Acoustic correlates of parental role and gender identity in the speech of expecting parents

Melanie Weirich, Adrian P. Simpson

Friedrich Schiller University Jena, Germany
melanie.weirich@uni-jena.de, adrian.simpson@uni-jena,de

Abstract

Differences between male and female speakers have been explained in terms of biological inevitabilities but also in terms of behavioral and socially motivated factors. The aim of this study is to investigate the latter by examining gender-specific variability within the same gender.

The speech of 29 German men and women - all of them expecting their first child but varying in the time they plan to stay at home during their child's first year (parental role) - is analyzed. Acoustic analyses comprise the vowel space size and the realization of the inter-sibilant contrast.

While the data is part of a larger longitudinal project investigating adult- and infant-directed speech during the infant's first year of life, this study concentrates on the recordings made before the birth of the child. Inter-speaker variability is investigated in relation to 1) the chosen parental role and 2) self-ascribed ratings on positive feminine attributes (gender identity).

Results show that both factors (planned duration of parental leave and the femininity ratings) contribute to the variability found between, but also within the same gender. In particular, the vowel space size was found to be positively correlated with self-ascribed femininity ratings in male speakers.

Index Terms: sex/gender-specific variability, social-indexicality, gender identity, acoustic vowel space, sibilant contrast

1. Introduction

From Speech Accommodation Theory (SAT, [1, 2]) we know that speakers vary in small phonetic detail depending on the communicative context, the situation, their interlocutor or their social identity. In addition to speakers expressing their social identity [3], listeners also interpret speech based on who they think the speaker is (e.g. regarding the speakers assumed age, gender or ethnicity) [4, 5]. Thus, speech conveys much more than the linguistic message, also comprising rich information about the speaker which then has to be interpreted by the listener. Whether the interpretation by the listener is equal to what the speaker intended is a different question. Some of these cues might be biologically inevitable, but some are definitely based on a (conscious or unconscious) choice of the speaker. The project adds to the growing body of research on fine-grained aspects of speech and social-indexicality, particularly concerning gender and sexuality [6, 7].

Research focusing on differences between male and female speakers has shown that these differences are also due to both biological and behavioral factors. We will refer to this variability therefore as sex/gender-specific variability, including both potential impact factors.

A typical difference between male and female speech is the acoustic vowel space size, which is on average larger in females than in males [8]. Both biological and behavioral reasons have been offered to account for this variability. While it is clear that physiological differences in vocal tract dimensions between men and women affect sex-specific differences in formants and vowel space size [9], cross-linguistic differences in the size of sex/gender-specific variability regarding vowel space size point to behavioral reasons as well [10, 11, 12].

Another particular salient cue of gendered speech is the spectral characteristics of /s/. Here too, both biological [13] and behavorial factors [14, 15] have been proposed to account for these sex/gender-specific differences.

While research on sex/gender-specific variability has mainly focused on investigating differences between male and female speakers, more recent studies have also highlighted gender-specific variability within the same gender. Here, self-ascribed gender *identity* has been found to play an important role ([14, 15]). One study [16] found that English speaking boys and girls from the age of four not only differ in their acoustics of /s/ but start to express their gender identity (measured by questionnaires completed by the parents) through their /s/-productions. By investigating parameters typically ascribed to the sex of the speaker within the same sex, differences between male and female speech independent of sex-specific biological inevitabilities can be detected.

Following this line of research, the present study particularly highlights intra-sex differences based on *gender identity*. In addition, the potential impact of *parental role* is examined. Gender identity is assessed directly by asking the participants for self-ascribed ratings regarding positive feminine attributes. Parental role is defined by the time participants plan to stay with their infant as primary caregiver during the child's first year.

In detail, the aim of this study is to analyze whether acoustic vowel space size and inter-sibilant contrast vary between speakers of the same sex depending on 1) chosen parental role and 2) self-ascribed ratings on positive feminine attributes.

2. Method

The data presented is part of an ongoing longitudinal project (*AVARE*) investigating *infant*- and *adult*-directed speech of 80 German and Swedish mothers and fathers. Recordings are made before the birth of the child and at several time points during the infant's first year. Participants differ in the time they plan to spend with their child (duration of parental leave).

http://dx.doi.org/10.21437/Interspeech.2017-1394

One of the aims of *AVARE* is to explore the potential impact of the time of the recording and the involvement in child-care on various acoustic aspects. In addition, socio-psychological factors including self-reported masculinity/femininity ratings are taken into account.

AVARE is still in its early stages, so here, we are presenting an analysis of adult-directed speech of 29 German speakers *before* the birth of the child. Analyses were made regarding the production of vowels and sibilants and their potential speaker-specific variation due to differences in 1) planned duration of parental leave and 2) self-ascribed positive feminine attributes.

2.1. Participants and speech material

The speakers (21-43 years) all live in Jena, Germany, and were expecting their first child. 14 females and 15 males took part in the analysis. All females (mothers = Mo) planned to take parental leave for at least 12 months. The males differed in the time they were planning to stay at home with the child as primary caregiver. Five speakers were planning to stay at home for at least three months when the child is between 5 and 10 months old (fathers = Fa), while the other 10 speakers were planning to work regularly during the child's first year (control fathers = CFa).

A picture description task directed to the experimenter was used to elicit several target words (animals and objects). Among them were the target words listed in Table 1 containing the vowels /i: ɛ a u:/ and the two sibilants /s, ʃ/ used for the present investigation. Table 1 gives an overview of the number of tokens analyzed.

Table 1: *Overview of the number of analyzed vowels and sibilants in the target words*

	i:	ɛ	a	u:	s	ʃ
target word	*Tiger Biene Wiese*	*Äpfel*	*Katze Tasse Wasser*	*Kuh Kuchen Buch*	*Tasse*	*Tasche*
	tiger bee meadow	apples	cat cup water	cow cake book	cup	bag
n	335	134	264	309	127	127

Fifteen different pictures were used (see Figure 1 for some examples). The number of tokens analyzed differs between the speakers due to differences in description style. However, each target word appeared at least three times on the pictures and for all further analyses at least three renditions of each vowel or sibilant could be used for all speakers.

Figure 1: *Some of the pictures used to elicit the target words*

2.2. Acoustic analyses

2.2.1. Vowel space

Formant frequencies (F1 and F2) were measured within the stable acoustic region in the middle of the manually segmented vowels /i: ɛ a u:/ part of the target words shown in Table 1. Mean formant values of each vowel category for each speaker were used to calculate the size of a speaker's acoustic vowel space. This vowel space was estimated by calculating the area of the polygon enclosed by the mean formant values of the vowels in the F2xF1 space (in Hz).

2.2.2. Sibilant contrast

Discrete Cosine Transformation (DCT) [17] was used to parameterize the shape of the sibilant spectra and, in particular, to quantify the acoustic contrast between /s/ and /ʃ/. DCT decomposes the signal into a set of half-cycle cosine waves and the resulting amplitudes of these cosine waves are the DCT coefficients (corresponding to the cepstral coefficients of a spectrum). Three DCT coefficients were used for the analysis. DCT1 is proportional to the linear slope of the spectrum, DCT2 corresponds to its curvature and DCT3 describes the amplitude of the higher frequencies. DCTs have been shown to provide a very effective separation between the four fricative types in Polish [18], and to be a reliable parameter to differentiate the very similar acoustic spectra of /ç/ and /ʃ/ in two dialectal regions in Germany [19].

For a quantification of the sibilant contrast, Euclidean Distances (EDs) in the DCT1xDCT2xDCT3 space were calculated for each speaker by pairing at least three repetitions of /s/ and /ʃ/ productions.

2.3. Personal Attributes Questionnaire

Self-reported gender identity ratings were gathered by means of the F+ scale of the *German version of Extended Personal Attributes Questionnaire* (GEPAQ-F+, [20]) which is a measure of socially desirable components of *feminine expressivity*. Participants were asked to rate various components, e.g. helpfulness, kindness, softness, etc. on a scale from 1 to 7. Then a mean score with high ratings reflecting high self-ascribed positive feminine attributes was calculated for each speaker and used for the correlations with the acoustic measures.

3. Results

3.1. Self-ascribed gender identity

Figure 2 reveals that both female and male speakers show variation in their self-ascribed femininity scores. Females show mean ratings between 4.7 and 6.6 with an average rating of 5.7; males vary between 3.6 and 5.9 and have an average score of 4.8. A Welch two sample t-test showed a significant difference between the genders ($t = 3.95$, $df = 25.562$, p-value < 0.001). In addition, the males show a larger variation with a bimodal distribution revealing two peaks around 4.5 and 5.5. Three of the five fathers that plan to stay at home (Fa, green in Figure 2) have rather high F+ ratings contributing significantly to the appearance of the higher peak.

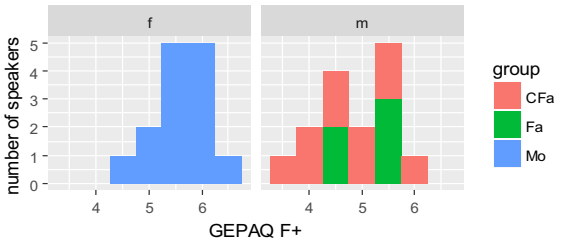

Figure 2: *Distribution of self-ascribed positive feminine attributes (GEPAQ F+, high numbers reflect high positive ratings) separated by gender and speaker group*

3.2. Vowel space, parental role and gender identity

Mean acoustic vowel space sizes were calculated for each speaker using the average values of F1 and F2 measured for /iː ɛ a uː/. Figure 3 shows the distribution of the estimated vowel space sizes (in kHz²) separated by speaker group, i.e. expectant mothers (Mo) and fathers varying in the time they plan to stay at home (CFa, Fa). As expected, the females show larger vowel space sizes then the males, interestingly, the fathers who plan to stay at home show vowel space sizes in between the control fathers (CFa) and the mothers (Mo).

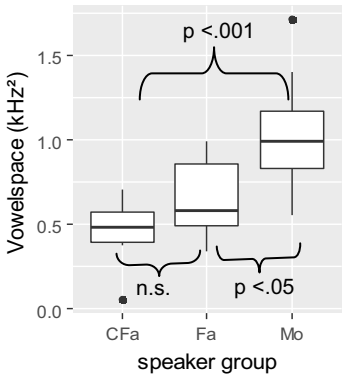

Figure 3: *Distribution of vowel space sizes separated by speaker group (mothers (Mo), control fathers (CFa), and fathers planning to stay at home (Fa))*

A linear model with vowel space as dependent variable and the speaker group as factor and a following post hoc test revealed significant differences between the mothers and the control fathers (Estimate: -0.57, $p < .001$) and also between the mothers and the fathers who plan to stay at home (though to a smaller extent, $p < .05$, Estimate: -0.38).

Figure 4 shows the vowel space size as a function of the self-ascribed positive femininity scores (by means of the GEPAQ F+ scale). Speaker groups are color-coded and individual speakers are plotted with their code numbers. As is apparent in Figure 2, females and males show variation in the GEPAQ F+ scores, with mothers (blue) having on average higher values than the control fathers (red) and the fathers who plan to stay at home having values overlapping both the expectant mothers and the control fathers.

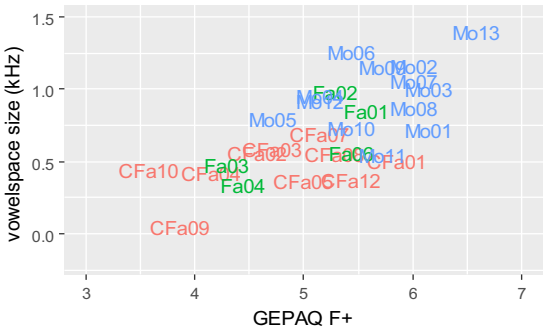

Figure 4: *Vowel space size as a function of gender identity (GEPAQ F+). Different speaker groups are marked by different colors (control fathers: red, fathers: green, mothers: blue), and individual speakers are plotted with their code numbers.*

To look for a potential relationship between vowel space size and gender identity (positive feminine attribute scores) that is independent of speaker gender, Pearson's product-moment correlations were run separately for each gender. While no significant correlation for the females was found, males revealed a significant positive correlation with $r = 0.54$ ($p < .05$). In other words, the higher male speakers scored on the feminine expressivity scale, the larger their vowel spaces were.

3.3. Inter-sibilant contrast, parental role and gender identity

To parameterize the acoustic contrast between /s/ and /ʃ/ EDs in DCT1xDCT2xDCT3 space were calculated between paired /s/-/ʃ/ tokens for each speaker. As mentioned above, at least three contrast pairs could be subjected to the analysis for each speaker (seven in one case). Figure 5 shows the distribution of acoustic sibilant contrast separated by speaker group. In line with the patterns found for vowel space size above, control fathers exhibit the smallest and mothers the largest contrast, while the fathers who plan to stay at home (Fa) have values in between the two.

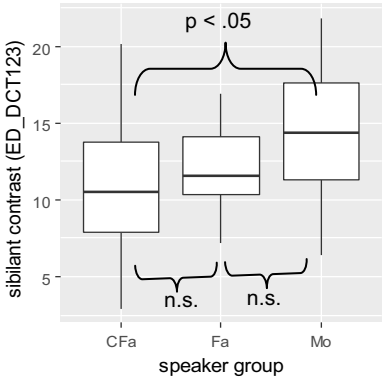

Figure 5: *Distribution of acoustic inter-sibilant contrast separated by speaker group (mothers (Mo), control fathers (CFa), and fathers planning to stay at home (Fa))*

A linear mixed model with ED as dependent variable, repetition as control variable, speaker group as test variable and the speaker as random effect showed a significant effect of speaker group (χ^2=8.9, df =2, p < .05). Post hoc test revealed that only the contrast between mothers and control fathers is significant (Estimate: -3.75, p < .05); while neither of the two father groups, nor the father (planning parental leave) and mother group differ significantly in the acoustic inter-sibilant contrast.

Pearson's product-moment correlations were run between sibilant contrast (mean values for the ED in DCT1xDCT2 xDCT3 space) and the GEPAQ F+ score. Again, separate correlation analyses were run for the two genders. As visual inspection of Figure 6 leads us to expect, no correlation was found within the female speakers. However, as was found in the vowel space analysis, for the male speakers a tendency for a positive correlation can be seen, which, however, marginally failed to show significance (r = .38, p = .07).

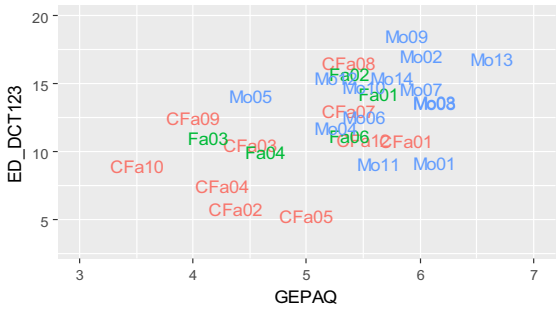

Figure 6: *Acoustic sibilant contrast (ED_DCT123) as a function of gender identity (GEPAQ F+). Different speaker groups are marked by different colors (control fathers: red, fathers: green, mothers: blue), and individual speakers are plotted with their code numbers.*

4. Discussion

This study investigated gender-specific acoustic variability in the speech of expectant mothers and fathers which could be related to differences in gender identity (measured by the GEPAQ F+ scale) and planned parental role (defined by the intention to take parental leave). We found that both factors contribute to the variation found in acoustic vowel space size and sibilant contrast. In detail, fathers, who plan to stay at home with their child on parental leave, show vowel space sizes and sibilant contrasts in between mothers and fathers who do not plan to stay at home. In addition, gender identity measures (self-ascribed femininity) correlate with the acoustic parameters in male speakers, with larger vowel spaces and sibilant contrasts in speakers with higher femininity ratings.

Of course, we are aware of the fact that the GEPAQ F+ score does not reflect all issues related to gender identity. Also, the planned parental leave does not mirror the actual part of the parent in the involvement in child care. Therefore, within the scope of the larger longitudinal project this study is part of, information about the involvement in child care is also being gathered by means of questionnaires at several time points during the child's first year. Together with acoustic analyses of several sex-specific phonetic features in speech, the project will show if the results presented here are sound. In a next step, listening experiments are planned to investigate the influence of varying vowel space sizes/sibilant contrasts on listeners' perception of masculinity/femininity.

5. Acknowledgements

This work is supported by a German Research Council Grant (WE 5757/2-1) awarded to the first author. We would like to thank all participating subjects.

6. References

[1] Giles, H. & Coupland, N. (1991) *Language: Contexts and consequences.* Keynes: University Press.

[2] Goldinger, S. (1998) Echoes of echoes? An episodic theory of lexical access. *Psych. Rev.* 105. 251–279.

[3] Foulkes, P. & Docherty, G. (2006) The social life of phonetics and phonology. *Journal of Phonetics* 34, 409438.

[4] Jannedy, S. & Weirich, M. (2014) Sound change in an urban setting: Category instability of the palatal fricative in Berlin. *Laboratory Phonology* 5(1), 91-122.

[5] Strand, E., Johnson, K. & M. d'Imperio (1999) Auditory-visual integration of talker gender in vowel perception, *Journal of Phonetics* 27, 359-384.

[6] Eckert, P. (2012) Three Waves of Variation Study: The Emergence of Meaning in the Study Of Sociolinguistic Variation. *Annual Review of Anthropology* (41) 87-100.

[7] Podesva, R. J. & J. Van Hofwegen (2016) /s/exuality in small-town California: Gender normativity and the acoustic realization of /s/. In Erez Levon and Ronald Beline Mendes (eds.) *Language, Sexuality, and Power: Studies in Intersectional Sociolinguistics*, Oxford: Oxford University Press, pp. 168-188.

[8] Hillenbrand, J., L. A. Getty, M. J. Clark, & K. Wheeler (1995) Acoustic characteristics of American English vowels, *J. Acoust. Soc. Am.* 97, 3099–3111.

[9] Fant, G. (1966). A note on vocal tract size factors and non-uniform F-pattern scaling, *STL QPSR* 4, 22–30.

[10] Bladon, R.A.W., Henton, C.G. & Pickering, J. B. (1984) Towards an auditory theory of speaker normalization. *Language & Communication* 4 (1), 59-69.

[11] Henton, C. G. (1995). Cross-language variation in the vowels of female and male speakers, in *Proceedings of the XIIIth International Congress of Phonetic Sciences*, Stockholm, Vol. 4, pp. 420–423.,

[12] Johnson, K. (2008) Speaker Normalization in Speech Perception, In D. B. Pisoni and R. E. Remez (eds.) *The Handbook of Speech Perception,* Blackwell Publishing Ltd, Oxford, UK, ch.15.

[13] Fuchs, S. & Toda, M. (2010) Do differences in male versus female /s/ reflect biological or sociophonetic factors?. In Susanne Fuchs, Martine Toda & Marzena Zygis (eds.), *Turbulent sounds. An interdisciplinary guide*, 281-302. Berlin: Mouton de Gruyter.

[14] Stuart-Smith, J. (2007) Empirical evidence for gendered speech production: /s/ in Glaswegian. In: Cole, J. and Hualde, J.I. eds. *Laboratory Phonology 9*. Series: Phonology and phonetics. Mouton de Gruyter, New York, USA, pp. 65-86.

[15] Munson, B. (2007) The Acoustic Correlates of Perceived Masculinity, Perceived Femininity, and Perceived Sexual Orientation. *Language and Speech* 50. 125-142.

[16] Li, F., Rendall, D., Vasey, P., Kinsman, M., Ward-Sutherland, A., & Diano, G. (2016) The development of sex/gender-specific /s/ and its relationship to gender identity in children and adolescents. *Journal of Phonetics*, 57, 59-70.

[17] Watson, C. I. & Harrington, J. (1999). Acoustic evidence for dynamic formant trajectories in Australian English vowels. *Journal of the Acoustical Society of America* 106, 458-468.

[18] Guzik, K. & Harrington, J. (2007). The quantification of place of articulation assimilation in electropalatographic data using the

similarity index (SI). *Advances in Speech Language Pathology* 9(1), 109-119.
[19] Jannedy, S. & Weirich, M. (2016). The Acoustics of Fricative Contrasts in Two German Dialects. in *PundP 2016 – 12. Phonetik und Phonologie Konferenz*, München.
[20] Runge TE, Frey D, Gollwitzer PM, Helmreich RL, Spence JT (1981) Masculine (instrumental) and feminine (expressive) traits. A comparison between students in the United States and West Germany. *J Cross-Cult Psychol* 12:142–162.