Interspeech 2024
1-5 September 2024, Kos, Greece

Gender and age based f_0-variation in the German *Plapper* Corpus

Melanie Weirich[1], *Daniel Duran*[2] *& Stefanie Jannedy*[2]

[1]Friedrich-Schiller-University, Jena
[2]Leibniz Center for General Linguistics, Berlin
melanie.weirich@uni-jena.de, duran@leibniz-zas.de, jannedy@leibniz-zas.de

Abstract

This study is the first exploration of data collected with the smartphone-app *Plapper* with which participants from Germany recorded themselves reading several sentences containing target sounds for future analyses of differences in fine phonetic detail and then donated their speech for inclusion to a large speech corpus. To this date, just short of 2.000 participants have contributed to this corpus. First analyses of differences in f0 on read speech from these German participants reveals an effect of age on mean f0 in females only and additional effects of self-rated femininity/masculinity (higher mean f0 in male and female speakers with higher self-rated femininity scores and vice versa). Also, there is an effect of region with speakers in the north of Germany inexplicably having lower mean f0 values than speakers from the other regions. Results leave room for speculation on the social meaning (what do speakers code and what do listeners interpret) of differences in f0.

Index Terms: big data, fundamental frequency, German, gender, femininity, age, socio-phonetics, *Plapper* corpus

1. Introduction

Language is not simply a tool for communication; it is also a vehicle for expressing social identities, affiliations, stances and power dynamics. Sociolinguistics examines the relationship between language and society, exploring how social factors such as *region* (in the sense of current location as opposed to birth place or place of socialization, which was also collected), *gender, age, sexuality, educational background* or *ethnicity* influence language use and variation. It investigates how individuals adapt their linguistic behavior to project social meaning and express and negotiate their identities within social groups. The study of social meaning is of high societal relevance: attitudes towards language, and towards individuals on account of their language use, are prominent in public discourse. Also, the study of social meaning strongly points towards a gap in the German research landscape and in the German language due to the long-standing tradition of studying regional variation in the sound domain rather than social variation. Investigating social meaning in fine phonetic detail in data collected from self-recordings of participants eliminates the observers' paradox. Aside from this major advantage, participants are not restricted to those living near our laboratory but can be recruited from all over the country and from differing social strata, therefore delivering a cross-sectional sample of speakers and their speech. Descriptive analyses of such widely varying data is the starting point for more deductive research testing very specific hypotheses on social variation reflecting identity negotiations.

In this study, we are looking at speakers' fundamental frequency (f_0) as the phonetic parameter and investigate f_0 variation in a diverse population sample on a wide range of socio-demographic scales. In particular, we will be analyzing f_0 variation as a function of *speaker gender*, *self-ascribed femininity/masculinity* ratings (on a scale), *sexual orientation* (on a scale), *age* and *region*.

F_0 is affected by both physiological and social factors. For example, many studies have investigated differences in f_0 patterns between male and female speech and it is known that differences in mean male and female f_0 are in part due to average differences in length and mass of the vocal folds [1]. However, cross-linguistic differences in the size and nature of gender-specific variation highlight the effect of cultural norms and expectations [2,3, for an overview see 4].

Also, the age of a speaker affects f_0: several studies have found a decrease of mean f_0 with increasing age in adult women [5-7], while results for adult males are less clear and studies report no changes, an increase and a decrease with increasing age [8-10]. Others point to a u-shaped change of mean f_0 with a drop in males over the age of 20 to 50 and an increase in older age [5-6]. The f_0 decrease in women and the f_0 increase in males with older age have been attributed to hormonal changes which affect vocal fold vibration [6, 11-12], while potential social factors on variation in f_0 across the life span are far from clear (but see 13 for the influence of becoming a parent on intra-individual variation in f_0 and vowel acoustics).

More generally, mean f_0 has been found to correlate with evolutionarily and socially relevant parameters such as hormone levels, physiology, physical strength but also perceptions of attractiveness, masculinity/femininity, and dominance [14–18].

1.1. Big data studies on variation in f_0

Two studies are particularly relevant to the present study since they also investigate the factors *gender* and *age* on variation in f_0 patterns and base their analysis on a large cross-sectional population sample.

The very recent and comprehensive study including more than 12.000 Icelanders (18-93 years, 56% females) [19] found that voice pitch is influenced by genetics but also varies between genders and across the life span in a systematic way. Female median f_0 measured in a reading task decreased by 0.8 Hz per year up until around 60 years of age (from 223.3 Hz at 20 to 30 years to 197.4 Hz at the age of 60). Male median f_0 was stable at 124.0 Hz up until around 60 years and increased by 0.6 Hz per year until reaching 136.5 Hz by the age of 90. Other factors adding to the variation explained were physiological traits for median f_0 (such as head mass, body fat percentage) and cognitive measures (such as verbal fluency) and the personality factor *openness* for variation in f_0.

The second large study particularly relevant for the present analysis is a study on German [20], the language also

10.21437/Interspeech.2024-1592

investigated here. It analyzes mean f_0 in more than 2.400 participants (40-79 years, 55% females), who were asked to count from 21 to 30. Interestingly, the study found mean f_0 to increase significantly with age for males but not for females. The difference to the Icelandic study might be based on the different age ranges investigated, with the German participants being older than 40 years thereby missing the onset of the decrease in mean f_0 in females. Mean f_0 was found to vary between four speaking voice intensity levels, with the one named "classroom voice" (third loudest) being most similar in values to the Icelandic study (f: 198 Hz. m: 130 Hz). However, mean f_0 measured in "conversational voice" (second loudest and considered to have the expected "normal" intensity during speaking) had surprisingly lower values for females than previously described (f: 169 Hz, m: 112 Hz). Socioeconomic status (SES) was found to be positively correlated with broader speaking voice ranges. Potential reasons suggested by the authors are that participants with higher SES may be more used to speaking in front of audiences and have more self-esteem, which helps to perform the requested task.

1.2. Gender

Both studies reported above look at the influence of gender as a categorical and two-dimensional concept. However, the traditional gender dichotomy is being called into question and there is consent that being a woman, or a man reflects sociocultural, learned, and biological factors [21]. However, when it comes to linguistic research most studies still assume that women and men are homogeneous groups – except for research focusing on differences within women and men based on sexual orientation of the speaker (e.g. [22-24]) or studies on transgender individuals (persons whose gender identity does not correspond with the sex assigned to them at birth, e.g. [25]). However, studies on cis-gender individuals (persons whose gender identity corresponds with the sex assigned to them at birth) also find much within-group variation in self-ascribed masculinity/femininity and gender role orientation [26-27]. Only recently, the inclusion of *gender* as a factor reflecting a multifaceted and continuous spectrum has caught on in research linking vocal cues with gender expression.

For example, [28] showed that heterosexual males index self-ascribed femininity through mean f_0 and vowel space sizes, and listeners are able to reliably detect this information using one-word stimuli. Similar results were obtained for females with correlations between self-ascribed and perceived femininity [29].

Cross-linguistic studies have shown that gender-specific variation is affected by cultural norms and expectations [2,13,29]. [13,29] found larger differences between genders in f_0 and vowel space size in German speakers than in Swedish speakers, due to lower values for Swedish females than for German females, accompanied by lower self-ascribed femininity ratings in Swedish females. Similar effects were found within a country: [31] conducted a cross-regional study on the effect of femininity on vowel acoustics comparing data from western Germany (former BRD) and eastern Germany (former GDR). These regions differ in gender role concepts due to different historical and societal developments. While the speakers in the east (with a more liberal gender role concept) indexed femininity in their vowel productions, speakers in the west (with especially males having more traditional gender role concepts) did not.

1.3. Aim of the present study

In this study we are analyzing data from a geo-spatially and socially diverse population on a wide range of socio-demographic scales. We analyze variation in f_0 considering the factors speaker *gender*, *self-ascribed femininity/masculinity* (on a scale), *sexual orientation* (on a scale), *age* and *region*.

2. Method

2.1.1. Smartphone app Plapper

We have developed the smart-device app *Plapper* to sample the German-speaking population in Germany: *Plapper* runs on Android and iOS and allows for people recording themselves in their own natural habitat without having to come to the lab thereby allowing data collection on a much greater scale outside of laboratories and addressing the observer's paradox. The smart-device app we are using is modeled after earlier versions of a similar app deployed in Switzerland [32], the UK [33-34] and Luxembourg [35]. Recordings are in wav-format and have a sampling frequency of 44.1 kHz. Data is continuously downloaded, and audio recordings are semi-automatically labeled using WebMaus [36] and specific segments are hand-corrected. Previous work has shown that the audio quality is suitable for analyses of spectral differences [32-35,37-39]. Participation is voluntary and completely anonymous, only meta data is collected via a questionnaire to be answered on a personal user profile upon opening the app.

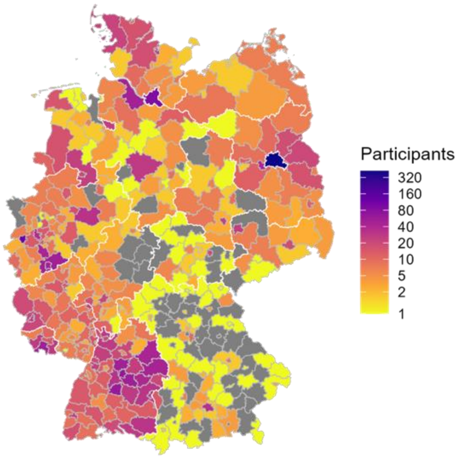

Figure 1: *Distribution of participants in Plapper corpus by district (grey = zero).*

2.1.2. Participants

This analysis is based on a data set consisting of 1.522 speakers from all over Germany and varying in educational background, profession, age, gender, and sexual orientation. Figure 1 shows the density distribution of the participants by districts.

Participants were asked which *gender* they feel they belong to (diverse, female, male, no answer) and rate their *sexual orientation* on the so-called Kinsey scale [40] (from 1 = exclusively heterosexual to 7 = exclusively homosexual) having also the opportunity to choose "no sexual contacts" or "no answer"). For the statistical analysis, participants with no sexual contacts (n = 38) and those identifying as diverse as their gender (n = 8) were excluded due to their currently rather small number compared to the other groups.

In addition, participants rated themselves on how masculine/feminine they feel on a masculinity-femininity spectrum (one-dimensional from 1 *very masculine* to 7 *very feminine*), also having the opportunity to choose "no answer".

Also, the users' present location (via the phone's geolocation information) and the birth year of the participants was collected. For the statistical analysis participants were grouped into three age groups reflecting important stages in the life cycle: younger (16-35), middle aged (36-55) and older (56-90). Table 1 gives an overview of the participants' average age, gender- and sexuality ratings (with standard deviation).

Table 1: *Summary statistics of participants.*

	F (n=932)	M (n=590)	D (n=8)
age	48.4 (14.6)	52.5 (15.6)	38.0 (19.5)
mf-spectrum	6.2 (0.93)	1.8 (0.9)	4.0 (0.93)
Kinsey-scale	1.6 (1.17)	1.6 (1.57)	2.6 (2.72)

2.1.3. *Speech material*

The data consists of self-recorded read speech via the smartphone app *Plapper*. Participants were asked to read a list of seven sentences, thereby resulting in a number of 9.441 (f: 5917, m: 3524) individual recordings and 1.522 speakers (f: 932, m: 590) included in the analysis.

2.1.4. *Acoustic analysis*

A custom *Praat* [41] (version 6.2) script was used to extract f_0 data from all audio recordings. An analysis was performed for each input file using the function "Sound: To Pitch (ac)..." with default parameters, except for "Very accurate" which was set to "on". From this analysis, the fundamental frequency in Hertz ("pitch") was extracted at each analysis frame (with a default time step of 0.01 seconds). The mean f_0 was then computed in *R* [42] for each recording (i.e. each read sentence), separately, ignoring missing values. To capture the variation in f_0, the inter-quantile range (IQR) between the 5%-percentile and the 95%-percentile was computed in *R* for each recording with the "quantile" function (using the default "Type 7" algorithm).

2.1.5. *Statistical analysis*

Linear mixed models were run in the *R* environment [42]. Test variables were added successively to the model. For ease of interpretation, categorical variables were treatment coded and the reference levels of the variables are given. P-values were estimated using model comparisons, i.e., by comparing models with and without the respective variables or interactions in question using likelihood ratio tests provided by the anova function (package *lme4*) [43]). For post hoc comparisons, the package *emmeans* was used [44]. Graphical explorations were done using *ggplot2* [45] and *ggeffects* [46].

As fixed factors we added *age group* (with the levels: 16-35, 36-55 and 56-90), *gender* (m, f), *gender spectrum* (ranging from 1 = very masculine to 7 = very feminine), *sexual orientation* (on a Kinsey scale from 1= exclusively heterosexual to 7 = exclusively gay/lesbian) and three factors representing the location of the speaker during recording: *LAT* (based on the latitude of the users' location with the levels *north*, *mid* and *south*), *LON* (based on the longitude of a users' location with the levels *west*, *mid*, *east*) and historical *region* (based on the federal state of the users' current location having the levels *old* (former west German BRD), *new* (former east German GDR) and *Berlin* (as a divided city after the second world war). *Participant* and *sentence* were added as random intercepts.

3. Results

3.1. Mean f_0

The final model which explains the data best regarding mean f_0 includes the interaction of *gender*age.group* ($\chi^2(2) = 48.8$, $p < .001$), and the fixed factors *gender spectrum* ($\chi^2(1) = 6.29$, $p < .05$), and *LAT* ($\chi^2(2) = 6.53$, $p < .05$) and Table 2 shows the summary statistics of the model. Post hoc tests reveal a significant decrease in mean f_0 in females only, with significant differences between the youngest and the middle aged group (Estimate: 15.31, $p < .0001$) and the youngest and the oldest group (Estimate: 17.62, $p < .0001$) but interestingly not between the middle aged and oldest group (Estimate 2.3, n.s.) pointing to a decrease of mean f_0 starting quite early in age (see also Figure 1).

Table 2: *Summary statistics of mean f_0 model (Number of obs: 9.441, groups: ID, 1.522; task_id, 7)*

Fixed effects	Estimate	Std. Error	df	t-value	Pr(>|t|)
Intercept	199.8	4.1	1241	49.23	< .001
genderM	-77.1	3.4	1524	-22.4	< .001
age.group (36-55)	-15.3	1.8	1511	-8.48	< .001
age.group (56-90)	-17.6	1.9	1507	-9.42	< .001
mf.spectrum	1.5	0.6	1512	2.6	< .05
LAT.mid	0.7	1.3	1509	0.5	0.6
LAT.north	-2.7	1.3	1503	-2.1	< .05
Gender:age group (36-55)	19.6	3.1	1549	6.4	< .001
Gender:age group (56-90)	19.7	2.9	1535	6.6	< .001

In addition, the effect of *gender spectrum* shows a link between self-rated masculinity/femininity scores and mean f_0: higher femininity is linked with higher mean f_0 and higher masculinity with lower mean f_0 - independent of gender (see Figure 3).

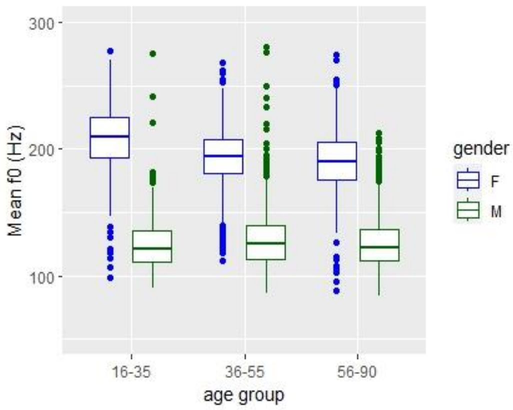

Figure 2: *Mean f_0 as a function of gender & age.*

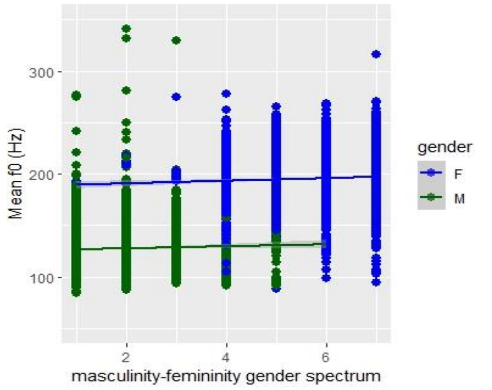

Figure 3: *Mean f_0 as a function of gender spectrum.*

Finally, while the historical *region* (former BRD vs. GDR) did not result in a significant effect (also not in an interaction with gender spectrum), *LAT* did: Post hoc comparisons show that speakers in the *north* have a lower mean f_0 than speakers in the *mid* of Germany (Estimate: -3.4, $p < .05$) independent of *gender* and *age* group) while the difference between *north* and *south* Germany fails to reach significance (Estimate: -2.7, $p = .09$).

3.2. Variation in f_0

For IQR we ended up with the final model including the interaction of *gender***age* group ($\chi^2(4) = 16.5$, $p < .01$) while all other factors did not lead to a better fit of the model. Post hoc comparisons show a significant difference between the youngest and the oldest age group, however with different directions between the genders: for females the IQR decreases with age (Estimate: -10.90, $p < .05$), for males it increases (Estimate: 17.37, $p < .01$).

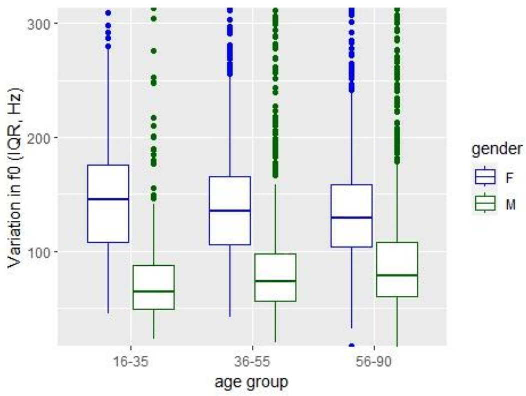

Figure 4: *Variation in f_0 as function of gender & age*

4. Discussion

We have been able to show that overall speaker gender as a non-binary variable affects mean f_0 in read speech [cf. 28]. Also, we have been able to show that f_0 lowering can be observed in younger age groups, long before hormonal changes related to menopause are supposed to have this effect on voice [cf. 19]. Males on the other hand are unaffected, mean f_0 remains stable across different age groups. Our results strongly suggest cultural mechanisms at work. With regard to variation in f_0, also gender specific changes across the life cycle have been detected with females showing smaller and males larger IQR with increased age. These patterns have not yet been reported in earlier studies [19-20]. In order to fully understand the factors playing out here, we will conduct more in-depth analyses of the data and meta-data in the corpus.

Although we also found an effect of region with speakers in the north of Germany having lower mean f0 values than speakers from the other regions, we need to be careful with interpreting this result. While the factor “region” reflects the current location of participants, it is not necessarily their origin. We also collected participants’ origin and location(s) where they were socialized. Further analyses will include ‘place of origin’ and ‘places of socialization’, leading us to a better understanding of the regional variation in mean f0.

The *Plapper* corpus [37-39] is growing steadily as more people participate in this crowd-sourcing effort. It is our firm belief that not only the number of speakers will significantly increase - potentially evening out to some degree the disparity in numbers between men and women. More importantly though, we are aiming at gaining greater regional coverage. As the map in Fig.1 shows, currently there are still districts from where we have no samples at all or where we are undersampled. Some of these areas may be very rural with a rather low overall population density, yet, still, it is our aim also to reach out to people in these places as one of our future goals is to do more small-scale analyses of speech patterns in remote places, comparing data from small towns, rural areas but also larger urban centers. The comparison of speech from different age groups allows for apparent time analyses, which can tell us much about differences in trajectories of sound changes which are just starting or which are already in progress.

A shortcoming of the corpus in its current state is that many participants have higher level academic degrees. Here, too, we will attempt to reach out more effectively to people with less formal education or from vocational professions as we are aiming for a cross-section of society. Therefore, at this point we cannot compare the effect of the level of formal education and how that affects speech and ultimately the implementation of fine phonetic detail. Moreover, Germany is rather unevenly populated. We plan to conduct more small-scale analyses comparing speech patterns of participants in different smaller cities or rural areas or with varying work-force compositions. In addition to read speech, we are currently also collecting spontaneous speech allowing for the analyses of speech directed towards a more formal (police) and a less formal (friend) addressee. The recording methodology allows for studying functional variation and differences in speech register.

Also, we envision analyses of data from within larger cities (i.e. more upscale vs. more blue collar areas or areas with more multi-lingual composition) and compare these across different urban regions. The assumption is that speakers in cities may be more progressive and diverse than speakers from less inhabited remote places. The *Plapper* corpus bears great potential filling the data gap that speech science undoubtedly must acknowledge: most of our models and theories are based on experimental data collected from rather homogenous and smaller speaker groups living near our laboratories and being accessible to us. With the onset of portable and distributed technological devices, this is bound to change now rapidly. More varied data becomes available now more rapidly allowing for more in depth analyses of differences between age groups, genders, social and geographical landscapes, speech tasks, levels of formalities etc. allowing to uncover the social mechanisms and factors contributing to differences in fine phonetic detail.

5. Acknowledgements

Work was carried out with funding by the DFG grant-# 416591334, to SFB1412/C02 to Jannedy & Weirich.

6. References

[1] Stevens, K. N. (1998). Acoustic Phonetics. MIT Press.

[2] van Bezooijen, R. (1995) Sociocultural aspects of pitch differences between Japanese and Dutch Women, *Lang Speech*, vol. 38, no. 3, pp. 253–265, 1995.

[3] Weirich, M., Simpson, A.P., Ojbro, J. & Ericsdotter Nordgren, C. (2019) The phonetics of gender in Swedish and German, Proceedings of *FONETIK* 2019, pp. 49–53.

[4] Simpson, A. P. & Weirich, M. (2020). Phonetic Correlates of Sex, Gender and Sexual Orientation. In: Mark Aronoff (Ed.) *Oxford Research Encyclopedia of Linguistics*, OUP, New York.

[5] Baken, R. J. (2005). The aged voice: a new hypothesis. *Journal of Voice*, 19 (3), pp. 317-325

[6] Linville, S. E. (1996). The sound of senescence. *Journal of Voice*, 10 (1996), pp. 190-200.

[7] Nishio, M. & Niimi, S. (2008). Changes in speaking fundamental frequency characteristics with aging. *Folia Phoniatrica et Logopaedica*, 60, pp. 120-127

[8] Verdonck-de Leeuw, L.M. & Mahieu, H. F. (2004). Vocal aging and the impact on daily life: a longitudinal study. *Journal of Voice*, 18 (2), pp. 193-202.

[9] Harnsberger, J.D., Shrivastav, R., Brown, W.W., Rothman, H. & Hollien, H. (2008) Speaking rate and fundamental frequency as speech cues to perceived age. *Journal of Voice*, 22 (1), pp. 58-69

[10] Decoster, W. & Debruyne, F. (2000). Longitudinal voice changes: facts and interpretation. *Journal of Voice*, 14, pp. 184-193.

[11] Abitbol, J., Abitbol, P. & Abitbol, B. (1999). Sex hormones and the female voice. *Journal of Voice*, 13, pp. 424-446

[12] Gugatschka, M., Kiesler, K., Obermayer-Pietsch, B., Schoekler, B., Schmid, C., Groselj-Strele, A., & Friedrich, G. (2010). Sex hormones and the elderly male voice. *J of Voice*, *24*(3), 369-373.

[13] Weirich, M., & Simpson, A. (2019). Effects of gender, parental role, and time on infant-and adult-directed read and spontaneous speech. *Journal of Speech, Language, and Hearing Research*, *62*(11), 4001-4014.

[14] Pisanski, K. & Bryant, G. A. (2019). The evolution of voice perception, in *The Oxford Handbook of Voice Studies* (Oxford Univ. Press, pp. 269–300 .

[15] K. Pisanski, A. Oleszkiewicz, J. Plachetka, M. Gmiterek, D. Reby, Voice pitch modulation in human mate choice. Proc. *Biol. Sci*. 285, 20181634 (2018).

[16] Aung, T. & Puts, D. (2020). Voice pitch: A window into the communication of social power. *Curr. Opin. Psych*. 33, 154–161.

[17] Puts, D. A. , Hill, A. K. , Bailey, D. H., Walker, R. S. , Rendall, D. Wheatley, J. R…& Ramos-Fernandez, G. (2016). Sexual selection on male vocal fundamental frequency in humans and other anthropoids. *Proc. Biol. Sci*. 283, 20152830.

[18] Apicella, C. L. , Feinberg, D. R. & Marlowe, F. W. (2007) Voice pitch predicts reproductive success in male hunter-gatherers. *Biol. Lett*. 3, 682–684.

[19] Gisladottir, R. S., Helgason, A., Halldorsson, B. V., Helgason, H., Borsky, M., Chien, Y. R. & Stefansson, K. (2023). Sequence variants affecting voice pitch in humans. *Science Advances*, *9*(23), eabq2969.

[20] Berg, M., Fuchs, M., Wirkner, K., Loeffler, M., Engel, C., & Berger, T. (2017). The speaking voice in the general population: normative data and associations to sociodemographic and lifestyle factors. *Journal of Voice*, 31(2), 257-e13.

[21] van Anders SM. (2015) Beyond sexual orientation: Integrating gender/sex and diverse sexualities via sexual configurations theory. *Arch Sex Behav*. 44: 1177–1213.

[22] Munson B, McDonald E, DeBoe N, White A. (2006). The acoustic and perceptual bases of judgments of women's and men's sexual orientation from read speech. *J Phon*. 2006; 34: 202–40.

[23] Sulpizio S, Fasoli F, Maass A, Paladino MP, Vespignani F, Eyssel F, Bentler D. (2015). The sound of voice: Voice-based categorization of speakers' sexual orientation within and across languages. *PLoS ONE*. 10: e0128882.

[24] Kachel S, Simpson AP, Steffens MC. (2018). Do I sound straight?: Acoustic correlates of actual and perceived sexual orientation and masculinity/femininity in men's speech. *J Speech Lang Hear Res*. 61: 1560–1578.

[25] Zimman, L. (2012). *Voices in transition: Testosterone, transmasculinity, and the gendered voice among female-to-male transgender people* (Dissertation, Univ. of Colorado at Boulder).

[26] Lippa RA. (2005). Sexual orientation and personality. *Annual Review of Sex Research*. 16: 119–153.

[27] Kachel S, Steffens MC, Niedlich C. (2016) Traditional masculinity and femininity: Validation of a new scale assessing gender roles. *Frontiers in Psychol*. 7: Article 956.

[28] Weirich M. & Simpson, A.P. (2018). Gender identity is indexed and perceived in speech. *PLoS ONE* 13(12).

[29] Weirich, M. & Simpson, A. P. (2019). Do you hear what I say? The expression and perception of femininity in female voices, Poster at the Workshop on *Social Meaning*, Leibniz-ZAS Berlin.

[30] Weirich, M. & Simpson, A.P. (2023). Changes in parents' phonetic parameters during the first year of a child's life. Comparative study of German and Swedish. In: *20th International Congress of the Phonetic Sciences (ICPhS),* Prague.

[31] Weirich, M. & Zahner-Ritter, K. (2023) Indexing femininity in vowel acoustics: A comparison between speakers in eastern and western parts of Germany. In: 20th *International Congress of the Phonetic Sciences* (ICPhS), Prague.

[32] Leemann, A. (2016) *Deutschklang*. Smartphone App.

[33] Leemann, A. (2017) *English Dialects*. Smartphone App.

[34] Leemann, A., Kolly, M.-J., Purves, R., Britain, D. und E. Glaser. 2016. Crowdsourcing Language Change with smartphone applications. *PLoS ONE* 11(1): e0143060.

[35] Gilles, P. (2019) Using crowd-sourced data to analyse the ongoing merger of [ɕ] and [ʃ] in Luxembourgish. Proceedings of the 19th *Int. Congress of Phonetic Sciences Melbourne*, Australia.

[36] Kisler, T. and Reichel U. D. and Schiel, F. (2017) Multilingual processing of speech via web services, *Computer Speech & Language*, Vol. 45, pp. 326–347.

[37] Jannedy, S., Weirich, M. & Leemann, A. (2018) The ecological validity of crowd-sourced data. Talk presented at *Phonetik & Phonologie in deutschsprachigen Länder*n 14, Univ. of Vienna.

[38] Jannedy, S. & Weirich, M. (2023) Plapper - a smartphone app for selfrecording: Exploration of vowel spaces. 5th *Phonetics & Phonology in Europe* (PaPE 2023), 157-158.

[39] Jannedy, S. (2022) Plappern für die Wissenschaft – Eine App gegen die Datenlücke. Bericht über das Forschungsjahr 2020/2021: *ZAS ; Impressionen*, 115-126.

[40] Kinsey, A. C., Pomeroy, W. B. & Martin, C. E. (1948). *Sexual Behavior in the Human Male*. Saunders, Philadelphia, PA.

[41] Boersma, P. & Weenink, D. (2022). '*Praat: doing phonetics by computer*, version 6.2.23. http://www.praat.org/'.

[42] R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[43] Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48.

[44] Lenth R (2022). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 853 1.8.2, https://CRAN.R-project.org/package=emmeans

[45] Wickham, H. (2016*). ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

[46] Lüdecke D. (2018a). ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. *J. of Open Source Software*, 3(26), 772.