VOCAL FOLD VIBRATORY PATTERNS IN BILINGUAL SPEAKERS OF LOW AND HIGH GERMAN

Jörg Peters, Marina Frank, Marina Rohloff

Carl von Ossietzky University Oldenburg
{joerg.peters, marina.frank, marina.rohloff1}@uni-oldenburg.de

ABSTRACT

The aim of this study was to investigate whether changes in stability, periodicity, and rate of vocal fold vibration could serve as indicators of cognitive load in bilingual speakers. We collected speech samples from 95 bilingual speakers of Low German (LG) and High German (HG), aged 15-88 years. Younger speakers had lower levels of jitter and shimmer, along with increased HNR, CPPs, and mean f0 while speaking LG, indicating a greater cognitive load compared to speaking HG. Conversely, older speakers showed minimal or no differences in performance between the two languages. The study found similar results when using the relative dominance of LG over HG instead of age as the predictor variable, which we calculated based on age of acquisition, frequency of use, and self-reported language proficiency. These results highlight the importance of considering vocal fold vibration as a potential marker of cognitive load during language use in bilingual speakers.

Keywords: Voice quality, cognitive load, bilingualism, language dominance, endangered languages.

1. INTRODUCTION

Research on task difficulty has identified several changes in vocal fold vibration that can serve as vocal cues to increased cognitive load [1]-[3]. Studies on speech tasks such as memory and Stroop tasks have found that increased cognitive load is associated with (i) reduced vocal fold perturbation, which results in more stable and synchronized vibratory patterns and is characterized by lower jitter and shimmer [4]-[8]; (ii) reduced additive noise in the voice source and increased periodicity of the speech signal, resulting in a higher Harmonics-to-Noise Ratio (HNR) and Cepstral Peak Prominence (CPP) [4][6][9][10]; and (iii) a higher rate of vocal fold vibration, or fundamental frequency (f0) [4][6]-[8][11][12].

The present study examines whether characteristics of vocal fold vibration can likewise indicate increased cognitive load in non-balanced bilinguals when speaking their weaker language (referred to as L2). Limited research exists on voice quality in bilingual speech, but there is some evidence that L2 speech tends to show a more stable vocal fold vibration and to have a higher mean f0. Decreased jitter and/or shimmer was observed in the L2 of Chinese-English and English-Chinese bilinguals [13][14]. Increased mean f0 was reported for a wide variety of bilingual language pairs including Brazilian Portuguese, Chinese, English, Finnish, French, German, Italian, Korean, Maori, Russian, and Spanish as L1 and/or L2 [14]-[24]. No conclusive findings are available for HNR and CPP in bilinguals [25].

However, many of these findings cannot be interpreted without reservation as evidence of increased cognitive load in the L2. First, it is not clear from many studies how well the L2 was mastered and how its use might have been associated with higher cognitive load. Second, it is not always clear whether the differences found in vocal fold vibration arose from higher cognitive load when speaking the L2 or from language-induced differences in voice quality [26][27]. Numerous studies on bilinguals have even been conducted for the purpose of finding language-dependent differences in voice quality, but only a few studies have controlled for language-specific influences by adopting a design in which the same languages have been acquired by bilinguals once as L1 and once as L2 [13][14][28]-[30].

The association of language use with cognitive load is of particular interest for the study of endangered languages, as increased cognitive load in speaking these languages can be expected to lead to a further decline in their use. Many endangered languages are no longer acquired monolingually but only as a non-dominant language alongside the majority language. This can make it impossible to establish a balanced design which accounts for language-specific influences by including the languages of interest both as L1 and L2. On the other hand, a comparison of voice quality in the use of endangered languages across generations could provide information about the cognitive load associated with speaking these languages without being masked by possible language-specific voice characteristics. For example, a decrease of vocal fold perturbation from older to younger generations in the L2 might indicate a cross-generational increase of cognitive load in speaking the L2, even if this language would generally be associated with less vocal fold perturbation than the L1. In the latter case, increased cognitive load in younger speakers might show up as an additive effect on vocal fold vibration.

A decline in language use from the older to the younger generation is reported for Low German (LG), which historically was spoken monolingually in northern Germany but now is only acquired alongside High German (HG). Recent surveys indicate a sharp decline in the number of speakers of LG and its frequency of use [31].

This study aimed to explore how using LG and HG affects cognitive load across bilingual speakers of different ages in a local speech community. We analyzed measures of vocal fold vibration that are known to be linked to cognitive load in speech production. From [31] we expected that for younger speakers, using LG would be more demanding than using HG, while for older speakers, using LG would be less or equally demanding compared to HG. Accordingly, we hypothesized that younger speakers would show a decrease in jitter and shimmer when speaking LG, as well as an increase in HNR, CPP, and mean f0. In contrast, we expected that older speakers would either exhibit no significant differences or demonstrate an opposite pattern of results for the two languages. Furthermore, we explored whether changes in voice quality could be more directly associated with the relative dominance of LG over HG in bilingual speakers. To accomplish this, we calculated a dominance score for LG and HG and used this score instead of age to predict voice quality features.

2. METHOD

2.1. Participants and data collection

We recruited 95 participants (47 females, 48 males, aged 15 to 88 years) from the municipality of Krummhörn in East Frisia, in the northwest of the federal state of Lower Saxony in Germany. After removing four participants with incomplete data, the final analysis included 91 subjects. All subjects were native speakers of the local variety of Low German, known as East Frisian Low German, and acquired High German simultaneously or at the latest when they entered elementary school.

To explore the relationship between age and language dominance, we administered a survey to all participants which included 55 questions adapted from [31] and [32]. These questions covered topics such as age of acquisition, frequency of language use, and language proficiency. Using the participants' responses, we calculated global language scores for LG and HG for each subject. To predict voice characteristics, we calculated a dominance score by subtracting the global language score for HG from that for LG. We used this score as an alternative predictor variable to the subject's age in our regression models. Positive values indicate dominance of LG over HG, while negative values indicate dominance of HG over LG. The dominance scores of our participants ranged from -119 to +58, within a possible value range of -174 to +174.

Speech samples were recorded with a head-mounted omnidirectional microphone (DPA 4066) and a portable digital recorder (Tascam DR-100 MKIII). The recordings were digitized at 48 kHz sampling rate with 24 bits/sample quantization.

2.2. Procedure, data selection, and acoustic analysis

We asked participants to recite the weekday names from memory in each language, assuming their familiarity with these names in both languages. To avoid order effects, we randomized the language order for each participant.

To minimize potential segmental influences, we selected the /a/ vowel of the last syllable of the LG and HG weekday names for *Monday*, *Tuesday*, *Thursday*, *Friday*, *Saturday*, and *Sunday* for analysis (N = 1092). In LG, these syllables correspond to *dag*, while in HG, they correspond to *tag*. In careful speech, LG *dag* is realized as [da:x] and HG *tag* as [ta:k]. However, in informal speech we also find the variants [da:x] and [dax] in both languages. Inspection of individual tokens revealed that LG onset consonants matched HG consonants in 42% of cases, while LG coda consonants matched HG consonants in 49% of cases. To minimize the impact of neighboring consonants, we extracted the steady-state portion of the target vowel. Furthermore, we excluded vowels that exhibited creak or other unusual vibratory patterns, and we eliminated all instances of the final list member, *Sunday*, from the analysis because they exhibited a high proportion of creak and a lower signal-to-noise ratio. To increase the reliability of measurements, we excluded vowels with a steady-state portion of less than 80 ms, leaving us with a total of 826 vowel tokens for analysis.

We obtained the five voice parameters local jitter, local shimmer, Harmonics-to-Noise-Ratio (HNR), Cepstral Peak Prominence smoothed (CPPs), and mean fundamental frequency (mean f0) from the extracted vowels using default settings in *Praat* [33]. For HNR, CPPs, and mean f0, recordings were downsampled to 16 kHz before analysis. Mean f0 was estimated with the cross-correlation algorithm.

2.3. Statistical analysis

For the dependent variables *Global Language Score* as well as *Jitter*, *Shimmer*, *HNR*, *CPPs*, and *Mean f0* we fitted generalized linear mixed models using the *glmmTMB* package [34] in R [35]. As fixed effects, we included the within-subject effect Language (LG vs. HG) and the between-subject factors Gender (female vs. male) and Age along with interactions between these factors. For all dependent variables, except *Global Language Score*, we fitted additional models with Dominance Score as a predictor variable in place of Age. As random effects, we had intercepts for both Subject and Item (weekday names), as well as by-subject random slopes for the effect of Language. Order of elicitation (LG or HG first) and consonantal context (*dag*, *tag*, etc.) did not enhance model fit and were therefore not included as random factors. Since the residuals for *Jitter*, *Shimmer*, and *Mean f0* were not normally distributed, we fitted generalized linear mixed models with the Gamma distribution and log-link function for these variables. For the remaining variables, we fitted generalized linear mixed models with the Gaussian distribution. Likelihood Ratio Tests were used to obtain p-values for the full models. We will focus on reporting the main effect of Language and interactions involving Language and either Age or Dominance Score. To examine contrasts between slopes for LG and HG, we utilized the *emtrends* function of the R package *emmeans* [36]. The level of significance was set at p = 0.05.

3. RESULTS

Figure 1 shows the global language scores for LG and HG as a function of Age. Statistical analysis revealed a significant main effect of Language (χ^2 = 55.5619, p < .0001), and a significant interaction between Age and Language (χ^2 = 28.7505, p < .0001), but no interaction with Gender. Contrast analysis showed a significant difference between the slopes for LG and HG, suggesting that in younger speakers, HG tended to dominate over LG, with this difference gradually diminishing with age (Estimate = 1.3337, SE = 0.2505, t-ratio = 5.324, p < .0001). Figure 1 indicates a reversal of the dominance ratio for both genders, with LG becoming more dominant than HG around the age of 75. Note, however, that our sample size was limited to only four male speakers aged over 75, while the oldest female speaker was 70 years old.

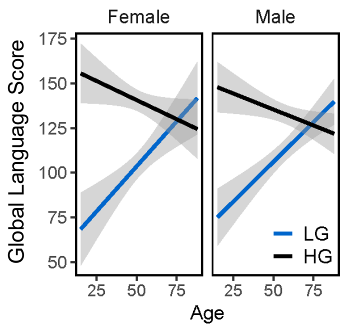

Figure 1: Regression lines for Low German (LG) and High German (HG) global language scores as a function of Age. Shading indicates 95% confidence intervals.

Figure 2 displays regression lines for *Jitter*, *Shimmer*, *HNR*, *CPPs*, and *Mean f0* for LG and HG, with Age (left panels) and Dominance Score (right panels) as predictor variables.

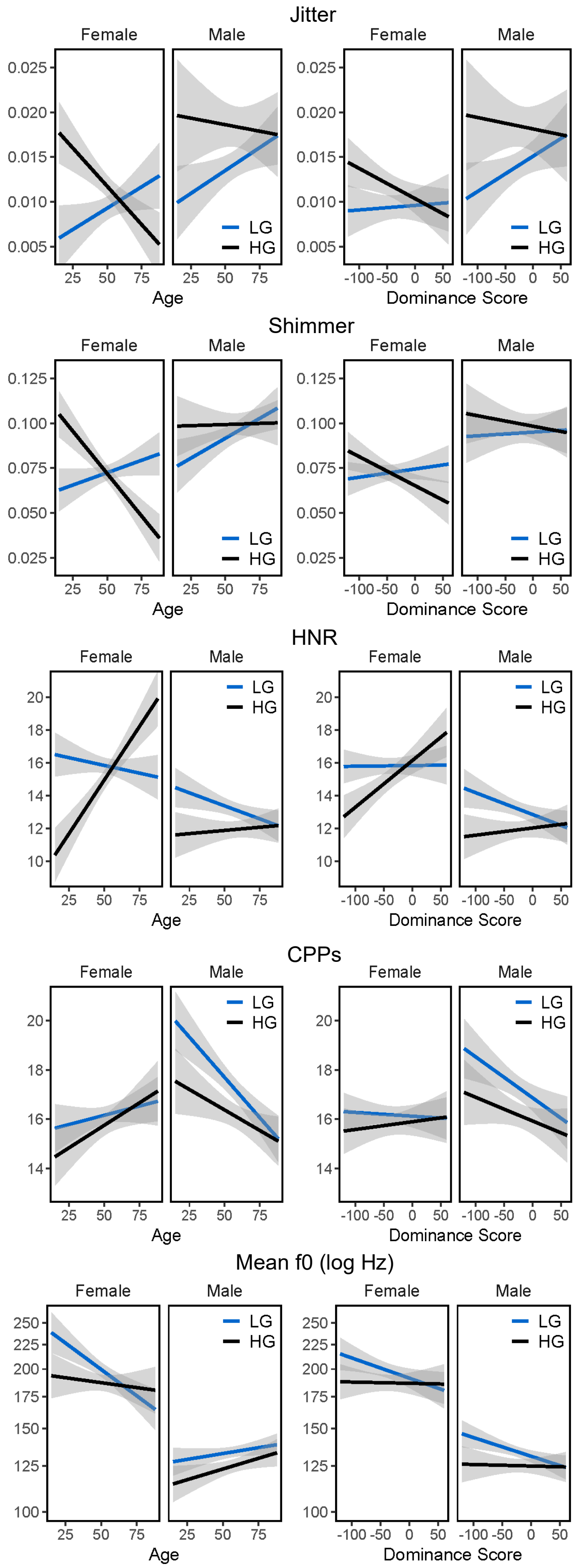

Figure 2: Regression lines for *Jitter*, *Shimmer*, *HNR*, *CPPs*, and *Mean f0* for LG and HG with Age (left panels) and Dominance Score (right panels) as predictor variables. Shading indicates 95% confidence intervals.

In the upper part of Table 1, we report significant interactions between Age and Language for *Jitter*, *Shimmer*, *HNR*, *CPPs*, and *Mean f0*. An additional interaction with Gender was found for *Jitter*, *Shimmer*, and *HNR*. The lower part of Table 1 presents significant interactions between Dominance Score and Language for all dependent variables, except for *CPPs*, where we observed only a main effect of Language.

Contrast analyses in Table 2 reveal in the upper part significant differences between the slopes of LG and HG for all dependent variables as a function of Age, with a greater difference between the slopes of LG and HG for *Jitter*, *Shimmer*, and *HNR* in female speakers compared to male speakers. In the lower part, we report significant differences between the slopes of LG and HG for *Jitter*, *Shimmer*, *HNR*, and *Mean f0* as a function of Dominance Score.

Table 1: Assessing model fit (χ^2) with Age (upper part) and Dominance Score (DS) (lower part) as predictor variables.

	Language	Age:Language	Age:Language:Gender
Jitter	31.5755***	23.5804***	4.0214*
Shimmer	0.2747	17.9844***	4.0175*
HNR	17.9314***	28.5918***	9.5736**
CPPs	15.5287***	4.4534*	0.0699
Mean f0	16.1781***	4.7376*	2.8276
	Language	**DS:Language**	**DS:Language:Gender**
Jitter	25.7868***	7.6465**	0.0015
Shimmer	0.2547	5.5552*	1.5447
HNR	16.4462***	12.7671***	0.3652
CPPs	14.9604***	1.2956	0.0031
Mean f0	15.5765***	5.4976*	0.0015

p-levels: * < 0.05, ** <0.01, *** <0.001

Table 2: Contrasts between LG and HG slopes for interactions involving Language and either Age (above) or DS (below).

Variable	Estimate	SE	*t/z*-ratio[1]	*p*
Jitter-female	0.0215	0.0047	4.584	<.0001
Jitter-male	0.0095	0.0037	2.566	.0103
Shimmer-female	0.0137	0.0032	4.215	<.0001
Shimmer-male	0.0053	0.0026	2.054	.0400
HNR-female	-0.1466	0.0256	-5.728	<.0001
HNR-male	-0.0462	0.0200	-2.312	.0210
CPPs	-0.0276	0.0139	-1.993	.0466
Mean f0	-0.0025	0.0010	-2.505	.0122
Jitter	0.0033	0.0012	2.767	.0057
Shimmer	0.0019	0.0008	2.304	.0212
HNR	-0.0019	0.0005	-3.557	.0004
Mean f0	-0.0009	0.0004	-2.345	.0190

[1]*t*-ratios: Gaussian distributions; *z*-ratios: Gamma distributions.

4. DISCUSSION

The aim of this study was to explore whether acoustic measures of vocal fold vibration reflect differences in cognitive load during bilingual speech. Specifically, we examined the acoustic parameters of /a/ vowels in weekday names produced by bilingual speakers in their two languages, LG and HG. We observed a decrease in jitter and shimmer and an increase in HNR, CPPs, and mean f0 in LG speech compared to HG speech in most participants. These results suggest that LG speech has higher stability, periodicity, and rate of vocal fold vibration, which has previously been associated with increased cognitive load [1]-[12]. However, we also observed that the effects on vocal fold vibration varied among speakers of different ages. For younger subjects, the acoustic analysis indicates higher cognitive load when speaking LG than when speaking HG. Interestingly, this difference reduced with increasing age. Therefore, our findings suggest that the additional cognitive load required for speaking LG instead of HG decreases as age increases. These findings are consistent with the strong correlation found between age and the dominance scores for LG and HG. Specifically, the younger the speakers were, the more HG dominated over LG. However, the dominance score of our speakers turned out to be a slightly weaker predictor for the dependent variables than age.

Our study demonstrates that acoustic measures can detect the effects of cognitive load on voice quality in bilinguals, irrespective of any language-specific effects. Although there may be differences in LG and HG speech with respect to the acoustic parameters examined, we were able to identify changes in vocal fold vibration associated with differences in cognitive load by comparing speakers differing in age and language dominance. However, our study was limited to reciting weekday names, and future research is necessary to investigate whether similar effects can be observed in other speech tasks.

5. CONCLUSION

Our findings support previous studies suggesting that increased stability, periodicity, and rate of vocal fold vibration result in decreased jitter and shimmer, as well as increased HNR, CPP(s), and mean f0, which can serve as indicators of increased cognitive load. Our results suggest that younger, non-balanced bilingual speakers of Low and High German exhibit these changes when speaking their weaker language, while the differences tend to disappear in older, more balanced bilinguals. These results align with surveys of language use and language proficiency, which indicate an increase of dominance of High German over Low German with decreasing age. By examining cognitive load through voice parameters, our study provides a valuable alternative to self-assessments and language proficiency tests in assessing the mastery of endangered languages.

6. ACKNOWLEDGMENTS

We gratefully acknowledge the support of the German Research Foundation (DFG) through grant no. PE 793/3-1.

7. REFERENCES

[1] Giddens, Ch., Barron, K., Byrd-Craven, J., Clark, K., Winter, A. 2013. Vocal indices of stress. A review. *Journal of Voice* 27, 390.e21-390.e29.

[2] Kirchhübel, Ch., Howard, D., Stedmon, A. 2011. Acoustic correlates of speech when under stress: research, methods and future directions. *Aviat. Space Environ. Med.* 18, 75–98.

[3] Puyvelde, M. van, Neyt, X., McGlone, F., Pattyn, N. 2018: Voice stress analysis: A new framework for voice and effort in human performance. *Front. Psychol.* 9, DOI: 10.3389/fpsyg.2018.01994.

[4] Boyer, St., Paubel, P.-V., Ruiz, R., El Yagoubi, R., Daurat, A. 2018. Human voice as a measure of mental load level. *JSLHR* 61, 2722–2734.

[5] Brenner, M., Shipp, Th., Doherty, E., Morrissey, P. 1985. Voice measures of psychological stress: Laboratory and field data. *Vocal fold physiology, biomechanics, acoustics, and phonatory control*, 239–248.

[6] Kappen, M., et al. 2022. Acoustic speech features in social comparison: how stress impacts the way you sound. *Sci. Rep.*, 12. DOI: 10.1038/s41598-022-26375-9

[7] Mendoza, E., Carballo, G. 1998. Acoustic analysis of induced vocal stress by means of cognitive workload tasks. *Journal of Voice* 12, 263–273.

[8] Rothkrantz, L., Wiggers, P., van Wees, J.-W., van Vark, R. 2004. Voice stress analysis. *International conference on text, speech and dialogue*, 449–456.

[9] MacPherson, M., Abur, D., Stepp, C. 2017. Acoustic measures of voice and physiologic measures of autonomic arousal during speech as function of cognitive load. *Journal of Voice* 31, 504.e1-504.e9.

[10] Yap, T., Epps, J., Ambikairajah, E., Choi, E. 2015. Voice source under cognitive load: Effects and classification. *Speech Communication* 72, 74–95.

[11] Griffin, G., Williams, C. 1987. The effects of different levels of task complexity on three vocal measures. *Aviat. Space Environ. Med.* 58, 1165–1170.

[12] Scherer, K., Grandjean, D., Johnstone, T., Klasmeyer, G., Bänziger, Th. 2002. Acoustic correlates of task load and stress. *7th ICSLP*, Denver, Colorado, USA, 1-4.

[13] Chong, Y. 2012. *Vocal characteristics of English and Mandarin produced by Mandarin-English and English-Mandarin bilingual speakers. A long-term average spectral analysis*. BA Thesis, Univ. of Hong Kong.

[14] Zhu, Sh., Chong, S., Chen, Y., Wang, T., Ng, M. 2022. Effect of language on voice quality - An acoustic study of bilingual speakers of Mandarin Chinese and English. *Folia phoniatrica et logopaedica* 2022, 1-10.

[15] Altenberg, E., Ferrand, C. 2006. Fundamental frequency in monolingual English, bilingual English/Russian, and bilingual English/Cantonese young adult women. *Journal of Voice* 20, 89–96.

[16] Chan, Y. 2010. *Acoustical differences in vocal characteristics between Cantonese and English produced by Cantonese-English bilingual adult speakers*. BA Thesis, University of Hong Kong.

[17] Mendez, N. 2017. *Do bilingual speakers shift fundamental frequency based on language acquisition or language dominance?* B.A. Thesis, Miami University.

[18] Mennen, I., Schaeffler, F., Docherty, G. 2007. Pitching it differently. A comparison of the pitch ranges of German and English speakers. *16th ICPhS*, 1769–1772.

[19] Ng, M., Chen, Y., Chan, E. 2012. Differences in vocal characteristics between Cantonese and English produced by proficient Cantonese-English bilingual speakers – A long-term average spectral analysis. *Journal of Voice 26*, e171-e176.

[20] Oh, H. 2011. Voice quality processing strategy of Korean learners of Chinese. *17th ICPhS*, 1526–1529.

[21] Peters, J., Frank, M., Rohloff, M. 2020. Pitch range variation in High German (L1) and Low German (L2). *Speech Prosody 2020*, 650–654.

[22] Ruß, A. 2015. *Stimmqualität zwischen Mutter- und Fremdsprache*. Saarbrücken: AV Akademikerverlag.

[23] Schwab, S., Goldman, J.-Ph. 2016. Do speakers show different F0 when they speak in different languages? The case of English, French and German. *Speech Prosody 2016.*

[24] Xue, St., Hagstrom, F., Hao, J. 2002. Speaking fundamental frequency characteristics of young and elderly bilingual Chinese' English speakers. A functional system approach. *Asia Pacific Journal of Speech, Language and Hearing* 7, 55–62.

[25] Engelbert, A. 2014. Cross-Linguistic Effects on Voice Quality: A Study on Brazilians' Production of Portuguese and English. *Proc. of the Int. Symp. on the Acquisition of Second Language Speech. Concordia Working Papers in Applied Linguistics*, 5, 157–170.

[26] Andreeva, B., Dimitrova, S. 2022. The influence of L1 prosody on Bulgarian-accented German and English. *Speech Prosody 2022*, 764–768.

[27] Kainada, E., Lengeris, A. 2015. Native language influences on the production of second-language prosody. *J. Int. Phon. Assoc.* 45, 269–287.

[28] Järvinen, K., Laukkanen, A.-M., Aaltonen, O. 2013. Speaking a foreign language and its effect on F0. *Logopedics, Phoniatrics, Vocology* 38, 47–51.

[29] Järvinen, K., Laukkanen, A.-M. 2015. Vocal loading in speaking a foreign language. *Folia Phoniatrica et Logopaedica* 67, 1–7.

[30] Theelen, M. 2016. *Adaptation of fundamental frequency in Dutch and English with second language effects*. B.A. Thesis. Univ. of Amsterdam.

[31] Adler, A., Ehlers, Ch., Goltz, R., Kleene, A., Plewnia, A. 2016. *Status und Gebrauch des Niederdeutschen 2016*. Mannheim: Institut für Deutsche Sprache.

[32] Poarch, G.J., Vanhove, J., Berthele, R. 2019. The effect of bidialectalism on executive function. *International Journal of Bilingualism* 23, 612–628.

[33] Boersma, P., Weenink, D. 2022. Praat: doing phonetics by computer [Computer program]. V. 6.2.16, retrieved 18 Aug. 2022 from http://www.praat.org/

[34] Brooks M., Kristensen K., van Benthem K., Magnusson A., Berg C., Nielsen A., Skaug H., Maechler M., Bolker, B. 2017. glmmTMB balances speed and flexibility among packages for zero-inflated Generalized Linear Mixed Modeling. *The R Journal* 9, 378–400.

[35] R Core Team 2022. R: A language and environment for statistical computing. R Foundation for Statist. Comp., Vienna, Austria, v. 4.2.1, www.R-project.org

[36] Lenth, R., Singmann, H., Love, J., Buerkner, P., Herve, M. 2019. emmeans: Estimated marginal means, aka least-squares means (R package v. 1.1)