Articulation of the STRUT vowel in L2 English speakers of Swiss German

Christine Gräppi, Carina Steiner and Adrian Leemann (Berne)

Abstract

Given the importance of pronunciation in effective communication, this study investigates Swiss German speakers' realisation of the English STRUT vowel. Anecdotally, such speakers use this sound more as an open-back vowel. The present study examined 54 students from two different school backgrounds reading out loud. The study involved a comparison of the students' vowel production with the standard British reference vowels as presented in school textbooks. The findings revealed that male Swiss German speakers produced the STRUT vowel closer to the target than their female counterparts. These results are surprising given that previous research indicated females attain more target-like articulations. In addition, the two school types present a significant effect on the target pronunciation.

1 Introduction

Numerous studies have been conducted in the field of second language acquisition (cf. Flege 1993; Gass 2013; Hansen Edwards/Zampini 2008; Munro/Derwing 1999; Pickering 2006) The investigation at hand encompasses a focus on pronunciation, as it bears relevance to the educational framework for primary and secondary education in Switzerland, specifically the *Lehrplan21* (primary education). This curriculum, spanning compulsory schooling from years 1 to 9, plays a pivotal role in guiding lesson planning and fostering collaboration among educators. Notably, within the *Lehrplan21*, a dedicated chapter addresses the critical aspects of English pronunciation and intonation in spoken discourse, aiming to ensure intelligibility. In addition, Swiss German schools teach with British English based course books, such as *Insight*, *Headway*, *New File* or Cambridge preparation exam course books. Besides, this chapter identifies specific linguistic features that pose challenges in production, such for example the differentiation between voiceless and voiced dental fricatives [θ] and [ð], or the voiced labiodental fricative [v], which are absent from the Swiss German phonetic inventory. Conversely, the curriculum for higher education merely stipulates the requirement for students to possess an adequate level of pronunciation proficiency, without delving into explicit details or specifications.

Intelligibility and accentedness have been the subject of extensive research. Questions arise regarding how the learner's first language influences their second language acquisition (cf. Gass 2013; Hansen Edwards/Zampini 2008; Munro/Derwing 1999). Previous research has shown that the realisation of consonant sounds can significantly impact intelligibility. Misunderstandings occur when speakers replace unfamiliar segments with similar sounds from their native

phonemic inventory (cf. Hansen Edwards/Zampini 2008). Prominent examples include various realisations of dental fricatives for first languages such as German, Dutch or Swiss German, often produced as /f/ or /s/ for voiceless dental fricatives and /d/ or /t/ for voiced dental fricatives by Swiss German speakers. The substitutions further range from [d], [d̪], [f], [tʰ], [s̺], [s] to even [ɾ] for the voiced and voiceless target sound. Dividing it into voiced and voiceless dental fricatives, voiceless are more often replaced by [f], [tʰ] or [s], and the voiced most frequently with [d], [d̪], [f] or [s̺] and sometimes with [ɾ], to name a few (cf. Graeppi/ Leemann 2019; Hanulíková/Weber 2010; Swan 2001; Wester/Gilbers/Lowie 2007; Zhao 2010).

Previous studies into Swiss German learners of English have analysed the grammatical structure of Swiss English and its language contact (cf. Durham 2007; Rosenberger 2009). Swiss German speakers have been found to struggle with specific consonants and vowels that do not exist in the Swiss German phonemic inventory. This results in significant differences in the realisation of vowels compared to Standard British English or General American English (cf. Graeppi/Leemann 2019; Leemann 2008). One vowel that poses significant difficulty for Swiss German speakers is the unrounded open-mid back vowel STRUT, which does not exist in their phonemic inventory. Research has argued that Swiss German speakers tend to produce the STRUT vowel as an unrounded open-back vowel such as /ɑ/ as in the General American English word *lot* (cf. Swan 2001). This differs significantly from its standard realisation in British and American English (cf. Leemann 2008). Moreover, it has been suggested that a speaker's realisation of a vowel might not be produced the way they intended to due to lack of time or effort, or the position of stress in a segment (cf. Thomas 2011). In this paper, the standard language is viewed as an abstract "set of norms to which actual usage may conform to a greater or lesser extent" (Milroy/Milroy 2012: 19). This underscores the importance of comparing the realisation of the STRUT vowel to standard variation as a guideline. Due to the fact that Swiss schools focus on British course books, this is the base to which the sounds are being compared.

The variation in the realisation of the STRUT vowel is not only of interest in SLA but also for variationist studies on English (cf. Clopper/Pisoni/de Jong 2005; Drummond 2010; Fabricius 2007; Ferragne/Pellegrino 2010). Thereby, the differences in vowel quality are fundamental to distinguish the various accents and Englishes overall (cf. Cunningham 2008). Variability has also been instrumental in our understanding of the realisation of the STRUT vowel in Swiss German speakers. In a sociolinguistic field of research, it has been shown that female speakers are more likely to produce a more standard form of a sound than their male counterparts (cf. Milroy/Milroy 1991; Trudgill 1972).

This present study aims to investigate how Swiss German speakers pronounce the STRUT vowel in comparison to the reference vowels and whether there is a difference in realisation a) between high school (Gymnasium) and vocational school (KV) students, b) male and female speakers and c) if a significant difference is visible between school years while taking exposure to the English language, anxiety and the importance of the speaker's own goals into account. The expectation is that, on the one hand, female speakers will have a more target-like realisation of the STRUT vowel overall. In terms of school types, we assume that the Gymnasium group outperforms the KV group due to their school career and the more in-depth learning of English in their school environment. Further, we expect a significant decline in the deviation from the

target vowel across the three school years from the first year to the third year of the Gymnasium or vocational school.

2   Methods

2.1   Speakers

Fifty-four students from two different school types (i. e., high schools and vocational schools) in Thun, Bern, and Lyss participated in this study. All participants were Swiss German natives (n=54), with six speakers having proficiency in a fourth language, after learning French and German in school as well. The additional languages (heritage language) spoken by a few of the participants were Albanian (n=2), Serbian (n=2), and Tamil (n=2). Prior to the interview process, explicit written informed consent was obtained from the participants themselves or their legal guardians, thus ensuring their voluntary and informed agreement to participate in the study.

2.2   Material

The students read two short texts and two lists of words including the target vowels. The 24 sentences contained in the texts comprised 23 target items excluding function words such as *but* and the verb *to be* in the past participle form *was* due to their weak form in connected speech (Table 1). Wells' lexical set and a list of words with the following context "/hVd/", were consulted for the extracted words (cf. Ladefoged/Johnson 2011; Wells 1982). The 64 reference vowels were taken from coursebook audios commonly used in Swiss schools such as *Headway*, *Life* and *Cambridge First Exam Preparation*. These audios were part of pronunciation as well as listening tasks. Notably, the coursebooks in the Swiss German curriculum are based on British English. In addition, most English teachers in this study follow British English variations such as Standard British or Scottish English, only one teacher spoke with an American accent. Nevertheless, the teachers did not take part in the study to be able to compare them to the students' realisations of the respective sound.

Extracted word list	One-syllable words	Two and three syllable words
strut	once	sometime
hud	one	another
	rushed	somebody
	up	running
	some	someone
	sun	everyone
	just	unusual
	such	sudden
		nothing

Table 1: List of target words in-text and word list

2.3   Procedure

Given the COVID-19 pandemic, 25 interviews were conducted via Zoom. Participants were asked to record themselves using their smartphones (14 iPhones, 11 Android) with their default recording app. Previous research suggests that their recording devices may affect the measurement of vowel quality, especially F2 in high-front vowels (cf. Ge/Xiong/Mok 2021; Sanker et al. 2021). The recordings were conducted in the participants' rooms at home, which were fully furnished. Another group of participants (n=29) was analysed and recorded in small, furnished, and sound-attenuated classrooms using a Zoom H2n mobile digital recorder. After the recorded session, metadata on the participants' educational background and English learning motivation was conducted. This was taken from Dörnyei (2009) and was collected via an online questionnaire. Thus, statements about the ideal L2 self, parental encouragement, instrumentality promotion, linguistic self-confidence, attitudes towards learning English etc. were included to receive further information on their language learning. Following the interview session, the set of target items was extracted for the analysis. We used Praat (Figure 1) to measure the three formants F1, F2 and F3 (cf. Boersma/Weenink 1992). Instead of measuring F1, F2, and F3 within the vowel onset to offset, we measured a short interval in the vowels' steady states, i. e., where their formant trajectories were most constant near or in the centre of the vowel.

The data was normalised to control for differences in vocal tract length to be able to allow a more accurate cross-speaker comparison (cf. Thomas 2011). Females have a shorter vocal tract, leading to higher formant frequencies for vowels than men. By using a normalisation method, exactly those physiological and anatomical differences can be neutralised so that a comparison is possible between the various speakers. The data was normalised through the online tool NORM (cf. Thomas/Tyler 2007). The vowel-intrinsic Bark Difference Metric was chosen for normalisation (cf. Flynn 2011) due to only one vowel being measured. Thereby, formants were converted into Z scores where the deviation of each score from the mean is calculated to then compare to the standard deviation. $Z_3 - Z_2$ is used to plot the front-back dimension of a vowel and $Z_3 - Z_1$ to plot the height dimension of the normalised vowels using the following formula (cf. Thomas/Kendall 2007):

$$Z_i = 26.81 / (1+1960/F_i) - 0.53$$

To test effects on target sound deviation, a linear mixed-effects modelling was conducted in R using the package lme4 (cf. R Core Team 2022). Regarding the outcome variable, deviations from the median reference value for z3z1 (F1) and z3z2 (F2) were calculated based on different sets according to gender. Item, gender, school year, school type, ideal L2 self, English exposure, and anxiety were entered as fixed effects, and random intercepts were entered for speaker and item. Thereby ideal L2 self is a concept that is based on the speaker's own goals and the extent to which they want to learn a language (cf. Dörnyei 2009). The feature ideal L2 self includes six statements that were measured according to the speakers' answers between (1): strongly disagree and (6): strongly agree. The statements mentioned the following:

- I can imagine myself living abroad and having a discussion in English.
- Whenever I think of my future career, I imagine myself using English.
- I can imagine a situation where I am speaking English with foreigners.

- I can imagine myself speaking English as if I were a native speaker.
- I can imagine myself as someone who is able to speak English.
- The things I want to do in the future require me to use English.

Anxiety, on the other hand, seen as a psychological construct and "defined as the subjective feeling of tension, apprehension, nervousness, and worry associated with an arousal of the autonomic nervous system" (Spielberger 1983: 15). In addition, studies have shown that female speakers tend to have higher levels of anxiety than their male counterpart (cf. Aida 1994; Koul et al. 2009).

The feature anxiety includes six statements as well that were measured according to the speakers' answers between (1): strongly disagree and (6): strongly agree:

- I get nervous and confused when I am speaking in my English class.
- I would feel uneasy speaking English with a native speaker.
- If I met a native English speaker, I would feel nervous.
- I would get tense if a foreigner asked me for directions in English.

In addition, two questions focus on how true they were for the speaker choosing between (1): not at all and (2): very much.

- How afraid are you of sounding stupid in English because of the mistakes you make?
- How worried are you that other speakers of English would find your English strange?

In the end, interactions between fixed effects were tested for and retained if they reached p<.01. Therefore, these two measures were added to find significant correlations for further explanations.

3   Results

3.1   Results of statistical modelling

To demonstrate how the STRUT vowel is produced in different contexts, the interactions between target deviation and factors such as school types, gender and school year were analysed using a linear mixed-effect model. The following section discusses only the statistically significant results ($p < 0.001$ and $p < 0.01$) regarding deviations from the median reference value for the F1 bark difference (z3z1) and F2 bark difference (z3z2), as determined by the model.

Fixed effect	Est	Std. Error	t value	Pr(>|t|)
(Intercept)	0.96	0.37	2.64	<.01**
Deviation Formant F2	0.03	0.05	0.63	0.53
GENDER (reference point: Female)				
Male	-0.28	0.08	-3.45	<.001***
SCHOOL_YEAR Year 2 (reference point: SCHOOL_YEAR 1)	0.03	0.08	0.34	0.74
SCHOOL_YEAR Year 3 (reference point: SCHOOL_YEAR 1)	0.12	0.10	1.21	0.23

Fixed effect	Est	Std. Error	t value	Pr(>|t|)
SCHOOL_TYPE KV (reference point: SCHOOL_TYPE GYM)	0.06	0.08	0.80	0.43
IDEAL L2 SELF	0.03	0.06	0.57	0.57
EXPOSURE	0.00	0.06	0.05	0.96
ANXIETY	-0.03	0.04	-0.58	0.57
Deviation Formant F2: (reference point: Female) GENDER Male	0.26	0.06	4.34	<.001***
Deviation Formant F2: SCHOOL_TYPEKV (reference point: SCHOOL_TYPE GYM)	-0.25	0.06	-4.19	<.001***

*<0.05 **<0.01 ***<0.001
(est. = estimate; Std. Error = standard error)

Table 2: Output summary of the final model

Table 2 illustrates the output summary of the final model of a linear mixed-effects model (LMM): (TARGET_DEV ~ FORMANT + GENDER + SCHOOL_YEAR + SCHOOL_TYPE + IDEAL + EXPOSURE + ANXIETY + FORMANT: GENDER + FORMANT: SCHOOL_TYPE + (1 | SPEAKER) + (1 | ITEM). The TARGET DEV is the dependent variable, the measure of deviation from the standard vowel sound. FORMANT is seen as the interaction term, modelling how the effect of the formant changes with respect to gender and school type. GENDER, SCHOOL_YEAR, and SCHOOL_TYPE are seen as categorical variables representing institutional factors. IDEAL L2 SELF, EXPOSURE, and ANXIETY are further categorical variables that were tested.

Looking at Table 2, the intercept estimate is at 0.96 which is the expected value of the dependent variable TARGET_DEV when all the other predictors are at their reference levels. In addition, the fixed factor FORMANT has two values: F2 (which appears in the table), and F1, which is in the intercept, accounting for variability between speakers. Male speakers diverge less from the reference medians (estimate -0.28), however not in F2 values (Deviation Formant F2: Gender Male, estimate 0.26). While SCHOOL_TYPE is not significant in terms of TARGET_DEV (0.06; $p = 0.43$), it is significant in terms of the effect of Deviation Formant F2 (-0.25; $p < 0.001$). Specifically, in KV schools, the pupils diverge less from F2.

Regarding SCHOOL_YEAR Year 2 (0.03; $p = 0.74$), and Year 3 (0.12; $p = 0.23$) both are not statistically significant, meaning the school year does not significantly affect TARGET_DEV. SCHOOL_TYPE KV (0.06; $p = 0.43$). Thus, indicating that there is no meaningful difference in TARGET_DEV based on school type. In addition, IDEAL L2 SELF (0.03; $p = 0.57$), EXPOSURE (0.00; $p = 0.96$), and ANXIETY (-0.03; $p = 0.57$) None of these predictors are statistically significant, meaning they do not contribute substantially to predicting TARGET_DEV in this model.

3.2   Effect of School Type

When focusing on the formants concerning school types (Figure 2), KV and Gymnasium have a similar deviation regarding F1 (blue line, z3z1) while the deviation of F2 (pink line, z3z2) is bigger for the Gymnasium group. The statistical analysis conducted reveals that no statistically significant difference exists between the groups of Gymnasium and KV in relation to the F1 variable. However, in terms of the F2 variable, the group of KV suggests less deviation.
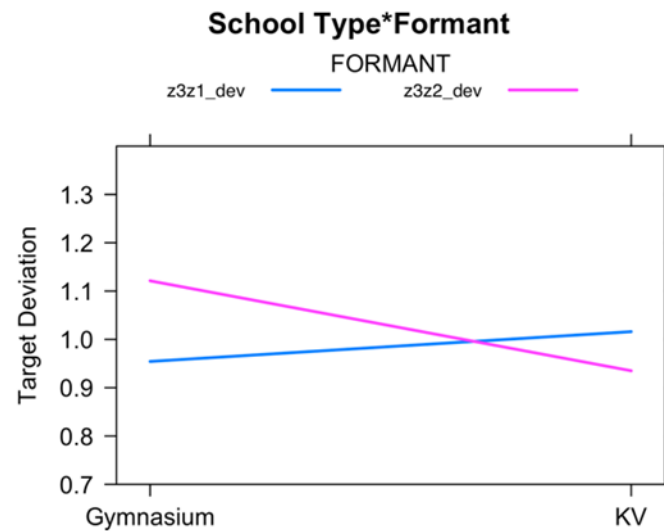

Figure 1: Interaction between school type and formant

While Figure 1 highlights the overall picture of the two groups, figure 2 provides the breakdown of the two school types compared to the reference vowels (pink). The two groups Gymnasium (green) and KV (yellow) present a strong overlap. The group Gymnasium appears to be more fronted and more open with their realisation of /ʌ/ than the group KV. The KV group seems to produce the target sound slightly more like the reference speakers in terms of the z3z2 (F2) variable and thus confirms the results in the model.
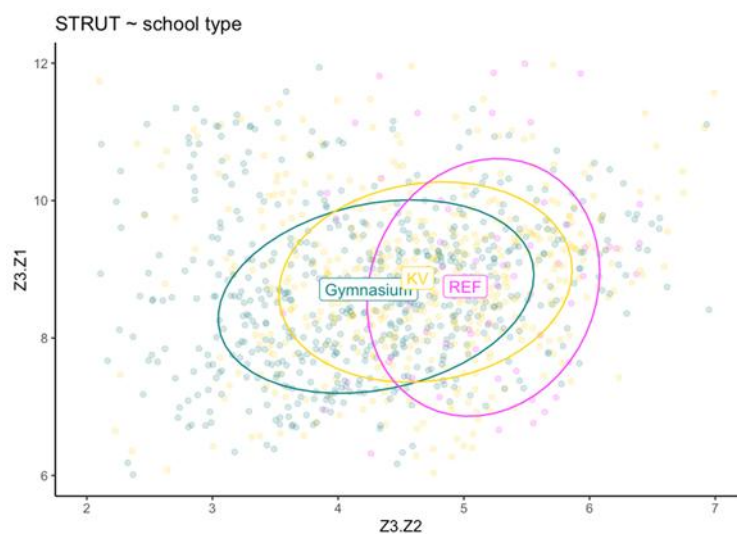

Figure 2: Realisation of STRUT within school types

3.3 Effect of gender

Figure 3 provides the relationship gender and formant. Thereby, the deviation of the F1 (z3z1) and F2 (z3z2) formants by female speakers compared to the reference vowel appear rather small, while male speakers present a wider deviation. Thus, the model shows that gender has a significant main effect on the outcome variable (TARGET_DEV), as do the interaction terms between Deviation Formant F2 and gender.
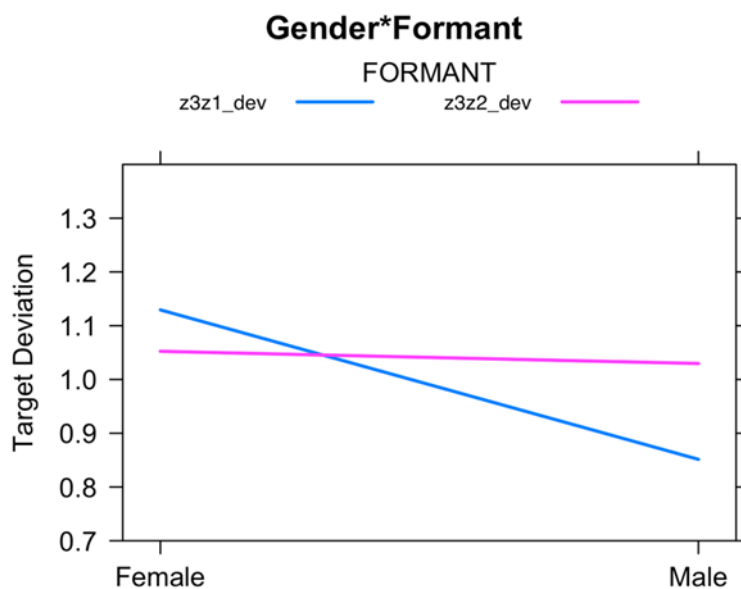

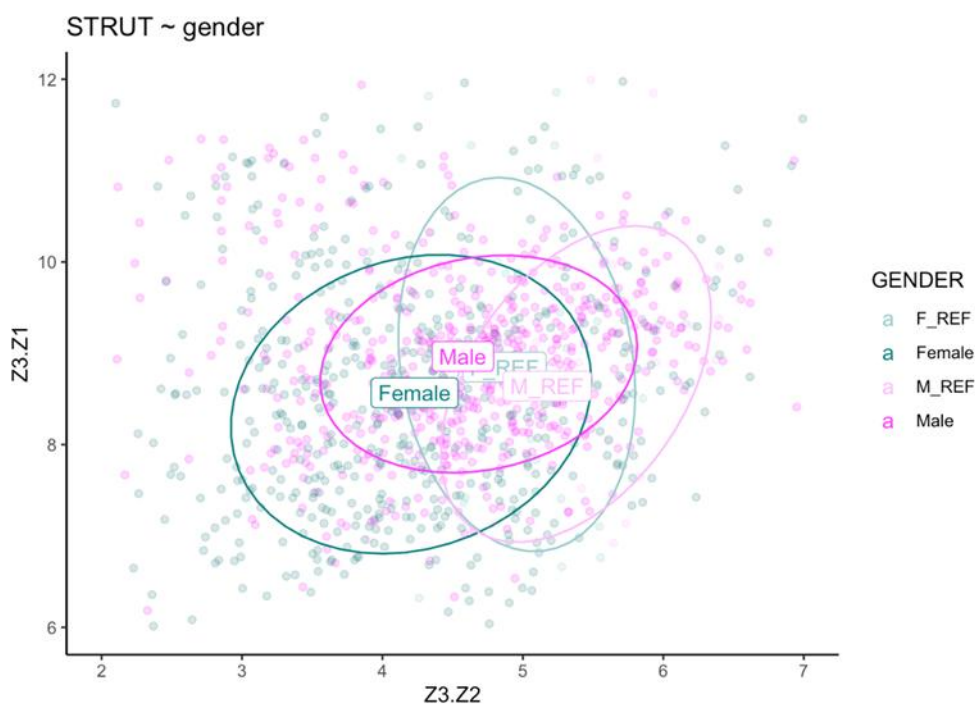**Figure 3: Interaction between gender and formant**

Figure 4: Realisation of STRUT between gender

From figure 4 above we can see that the male group provides a more concentrated picture while the female speakers are slightly more spread out in their pronunciation. According to this figure, the female speakers tend to be further away from their respective reference vowels, especially in terms of z3z1 (F1). The non-native male speakers, however, are closer to the target native female realisation and show a more condensed realisation.

While significant effects on school type and gender were found, no effect was present between school years. In addition, exposure to English or their Ideal L2 Self did not show any significant effects in terms of TARGET_DEV and thus is not further discussed.

4   Discussion

This study investigated the degree to which Swiss German speakers realise the STRUT vowel similar to the reference vowels of British English teaching coursebooks. We anticipated more deviation to occur for KV, due to the (lower) school level as well as for male speakers.

4.1   Overall realisation of STRUT

Overall, the results indicate that while there was no main effect for the formants, interactions between formant and school type as well as between formant and gender were identified. As for the Gymnasium and male speakers, F2 deviates more from the target sound than F1. One of the reasons for this result may be attributed to the difficulty of teaching the students how to move the tongue slightly backwards or forwards, as opposed to teaching the height of a vowel (cf. Fischer-Jørgensen 1985).

Moreover, it could be argued that while the STRUT vowel does not exist in the Swiss German phoneme inventory, L2 English speakers are more likely to assimilate the segment based on Flege's Speech Learning Model (SLM) (cf. Flege 1995; Flege/Bohn 2021). This model aimed to account for variation in the extent to which individuals learn – or fail to learn – to accurately produce and perceive phonetic segments in a second language. Explaining that phonetically identical or similar L2 sounds fit into the existing mental categories of L1, whereas new L2 sounds do not connect to those categories at all. If different L2 sounds are perceived as the same L1 category, it becomes challenging to distinguish between them (cf. Strange 2007). Therefore, it is essential to consider the processes by which individuals acquire L2 phonemes and how such L2 acquisition influences their production and perception of native language (L1) phonemes (cf. Flege/Bohn 2021). Flege's revised version of the SLM additionally emphasises a **phonetic factor**, where a new phonetic category for an L2 phoneme relies primarily on three factors: "(1) the sound's degree of perceived phonetic dissimilarity from the closest L1 sound, (2) the quantity and quality of L2 input obtained for the sound in meaningful conversations, and (3) the precision with which the closest L1 category is specified when L2 learning begins" (ibid.: 65). The first factor addresses the similarity or difference of an L2 sound compared to sounds in a learner's first language. A useful example to illustrate this factor is the replacement of the voiceless dental fricative /θ/ with [f], [θ], or [s]. In particular, [f] is phonetically close to [θ], which is predominantly used in Swiss German L2 English speakers (cf. Graeppi/Leemann 2019).

The second point emphasises the importance of exposure to the L2. The more varied and authentic the input is, such as in conversations or media, the better learners can understand and produce the new phonemes (cf. Flege/Bohn 2021). Thus, quality input that includes rich contexts helps learners grasp the nuances of how a sound is used in meaningful communication, reinforcing correct pronunciation. For this study, participants completed an extralinguistic questionnaire. One section of the questionnaire asked them to rate their current exposure to English in various contexts using a scale ranging from "not at all" to "very much". These contexts included interacting with friends and family, watching TV shows and movies on platforms like Netflix, listening to English radio and music, reading, and using language tapes or engaging in self-study. Around 77.7% replied with "quite a lot" and "very much" when watching TV or series, and 61.1% regarding listening to English radio/music. While 23.3% of the participants highlighted they are exposed to English "quite a lot" through friends, only 10% said they are exposed to it "very much". Thus, it shows that around three-quarters of the participants are regularly exposed to conversational English. It is worth noting that quite a few Swiss German participants mentioned they would often converse in English with each other during their leisure time. This practice likely enhances their conversational fluency in English. However, in these informal interactions, pronunciation may not receive as much attention or correction due to English being used as a lingua franca.

The third aspect, precision of the closest L1 category, focuses on the point when learners begin acquiring an L2, their existing L1 categories heavily influence their perception and production of L2 sounds (cf. ibid.). If they have a clear and precise understanding of the relevant L1 sound, they may be able to transfer that knowledge effectively. Understanding the correct articulation of specific sounds, particularly tongue placement, is crucial for accurate pronunciation (cf. Fischer-Jørgensen 1985; Ladefoged/Disnery Ferrari 2012). For example, knowing how to position the tongue correctly is essential for producing the STRUT vowel sound according to the rules. As mentioned at the beginning of the paper, the STUT vowel does not exist in Swiss German and thus makes it more difficult to produce in general.

Furthermore, an apparent feature observed in Figures 3 and 5 is the considerable variability in the realisation of the target phoneme. This serves as evidence for the distinctiveness of individual and inter-speaker variations in the articulation of this phoneme. The results of this study demonstrate that participants did not fully produce the STRUT vowel as per the reference given. As an incorporated form of the concept of assimilation, the term undershoot may be applicable in this context, as F2 is highly influenced by the place of articulation of neighbouring segments (cf. Thomas 2011). The term *undershoot* is explained in the context of this paper as a vowel that does not always attain the position that the speaker intends because of a lack of time or effort (cf. ibid.). Previous research into undershooting has shown that it is influenced not only by the rate of speech but also by where the stress was placed within a word and can be language- and speaker-specific (cf. Engstrand 1988; van Son/Pols 1990). Due to English being a stress-timed language, this could be a reasonable explanation for L2 learners of English to produce the sound accordingly. The term further incorporates assimilation, where one sound becomes more like a neighbouring sound, which was previously mentioned.

Another aspect that should be mentioned is the question of intelligibility due to its connection to second language pronunciation and the debate between various fields of research. On the one

hand, it had been argued that pronunciation needs to be native-like which stands in opposition to researchers on language change (cf. Munro/Derwing 1999). Beinhoff (2013: 35) highlights that "[n]on-native speaker accents in English are not necessarily less intelligible than native speaker accents in English". On the other hand, Flege (1993) once argued that pronunciation is essential for effective communication because speech carries both emotional and social meaning. Many people speak a foreign language with an accent and are understandably concerned about how their accent might affect their communication skills or shape others' opinions of them. Accents are a natural, inevitable, and significant outcome of language variation (cf. Walker 2022).

Vowels are particularly interesting to examine in this context due to their greater variation compared to consonants in English. The same word can be pronounced in numerous ways across different varieties of English. One prominent example is the STRUT vowel, as presented in this study, which exhibits differing qualities within England, as seen in the pronunciation of the word *but* – /bʌt/ in the southern regions versus /bʊt/ in the northern regions. Because of this variation, Jenner (1989) argued that native accents vary so widely in vowel quality that it is impossible to say these differences are crucial for mutual understanding. Similarly, there is no specific set of vowel sounds that a foreign learner must master, as different native accents do not include all the vowel contrasts found in Standard British English or any other standard variety. The problem that appears, however, concerns the length of the English vowel, which is vital in distinguishing minimal pairs such as in *sheep* vs *ship* due to the change in meaning. Thus, stating that it is more important to actually focus on the long-short differences between vowels rather than achieving any exact, native-speaker quality. According to research "vowels in open syllables have been shown to have approximately the same length as in syllables closed by a voiced (lenis) consonant. However, when an accented syllable is closed by a voiceless (fortis) consonant such as /p/, /t/, /k/ […], this has the effect of shortening the preceding vowel" (Walker 2022: 35).

Nevertheless, having seen these features the problem of miscommunication can still occur if a vowel sound is pronounced in different ways every time it is used in a sentence. Thus, vowel qualities for English language learner pronunciations need to be consistent in the same way to be able to be understood.

4.2 Differences between schools

When considering the difference in the realisation of STRUT between the two school types, the results indicate an overlap in the pronunciation of the vowel between the two groups. Nevertheless, the KV group appears to be closer to the reference vowel. This phenomenon can potentially be attributed to the fact that participants doing an apprenticeship are exposed to a greater degree of English language immersion within their professional milieu. This observation aligns with Ellis' theoretical proposition, which posits that second language proficiency is significantly influenced by social conditions and attitudes (cf. Ellis 2006). Furthermore, certain students reported utilising English-language tutorials as a means to enhance their comprehension of work-related topics. This was often due to the greater availability of relevant online resources in English as opposed to German. Thus, it seems that these students are more likely to be exposed to more English varieties or accents. In the case of Gymnasium students, it appears that

their involvement across multiple school subjects may diminish their singular focus on a particular one such as English in this case. One of the questions asked about how different factors affected the participants' English learning such as: "interacting with friends", "interacting with family", "reading", "language tape/self-study", "watching TV/Series or Netflix", "listening to English radio/music". Overall, 48.8% of the participants pointed out that their friends play a crucial role in learning the language, while 57.7% say reading helps a lot as well and around 70% agree that Netflix/watching TV would be very useful in their learning process. In addition, students mentioned in the interviews how much they enjoy different accents; in many discussions, the preferred accent was British English. Besides this, the students were asked whether they want to achieve a native-like pronunciation at some point which was answered "very much" by around 63% (Gymnasium students) and only 23.3% by the KV students (and the answer "quite a lot" 30% Gymnasium and 40% KV). When discussing the native-like accent, the variety of pronunciations need to be considered as well. This study, however, only compared the sounds to the Standard British/Standard American English as a base to be compared to. Thus, it is of importance that teaching English includes various accents and Englishes in the curriculum to highlight the plethora of vowel sounds that change depending on the region, as an example. Nevertheless, it depends on the learner themselves what they want to focus on when learning the language: some participants mentioned that they need English in a professional context where people have other native language backgrounds; thus, using English as a lingua franca. Others mentioned that they want to learn English to be able to travel when they are done with their apprenticeship or school. Thus, it presents how individual the interest of learning another language is (cf. ibid.).

Although this study included variables such as exposure and ideal L2 self in a questionnaire, no significant results were found. Further research is needed to explore the influence of motivation in this specific setting. In addition, the teacher whose students were participating highlighted that they did not really focus on pronunciation tasks even though they were present in the course books.

4.3   Differences between gender

Contrary to expectations, this study found a higher deviation between reference vowels and female speakers. Thus, the findings of the current study do not support previous research. In many studies, it has been argued that female speakers produce the more prestigious variation in language (cf. Milroy/Milroy 1991; Trudgill 1972). One of the reasons for the male speakers' more target-like realisation could be their interaction with native speaker groups through gaming or work-related. During the interviews, participants consistently demonstrated a heightened interest in English communication, whether it be in social interactions with friends or in customer interactions. Moreover, a greater number of male participants expressed a preference for English-language instructional videos over their German counterparts, highlighting the alignment of these findings with previous research that has summarised diverse learning styles (cf. Dong/Liu/Zhou 2023). Language learners often exhibit varying levels of success, with some performing well while others struggle. Every learner might have different learning strategies to succeed in language learning. Thus, language learning strategies (LLS) or style "are purposeful mental actions […] used by a learner to regulate [their] second or foreign language learning"

(Oxford 2018: 81). LLS exhibits a certain complexity as it is formed through many external factors such as task requirements, personal preferences, and educational policies (cf. Oxford 2018). Peacock/Ho (2003), for instance, observed in their study that female English learners showed better use of the memory and metacognitive strategies, such as reviewing lessons and practising. Multiple factors contribute to the process of second language acquisition, encompassing not only a variety of learning styles of the speakers but also individual personality traits, internal motivation, self-esteem, anxiety as well as willingness to communicate (WTC), to only name a few (cf. Dong/Liu/Zhou 2023; Ellis 2006; Heyde 1979; Mruk 2013; Wucherer/Reiterer 2018; Zarrinabadi/Pawlak 2021). WTC has been subject of discussion in second language research as it constitutes a crucial determinant of a learner's potential language proficiency. Besides WTC, motivation is another salient feature in language learning which includes the effort and desire to learn the language. Within this field of research, females are reported to outperform male L2 learners due to a higher level of motivation (cf. Kissau 2006). According to other researchers, female participants appeared to be more intrigued by learning another language, its culture, and the speakers of that language than their counterparts (cf. Dörnyei/Clément 2001; Mori/Gobel 2006). Further, it is argued that boys are more interested in learning a language to get essential qualifications or to have better career opportunities, girls, on the other hand, desire to learn a language merely for communicational reasons (cf. ibid.). This current data indicates minor differences between female and male speakers regarding their intrinsic interest in the English language, particularly in aspects such as its sound, rhythm, and vocabulary. Female participants exhibit a greater curiosity toward these linguistic elements compared to their male counterparts. However, when analysing the aspects of instrumental motivation, both groups appear to demonstrate the same perception of the language's necessity. This aligns with Francis (2000: 88) who argues that "increased ambition, coupled with a feeling that opportunities in the workplace are skewed against them, is what has provided girls with new motivation for achievement at school".

The findings of this present study pertaining to male speakers exhibiting a higher degree of target-like realisation in the STRUT vowel align with previous research indicating that males tend to excel in phonetic imitation, whereas females display greater proficiency in grammar (cf. Wucherer/Reiterer 2018). Furthermore, López Rúa (2006) even found reasons for the girl's better performance through "interaction of neurological, cognitive, affective, social and educational factors" (Alexious 2016: 88).

In contrast, some researchers have argued that female learners may be slower in acquiring a second language due to their role as "guardians of the home language and culture" (Piller/Pavlenko 2001: 4), which could hinder their willingness to adopt a new language. While male speakers appear to favour a more localised variation, which carries some kind of identity-based social meaning in the local community, whereas females identify more with supra-local variants in speech. Nevertheless, it is argued that studies on gender in connection with speech should be more "context-sensitive which treats gender as a system of social relations and discursive practices whose meaning varies across speech communities" (ibd.: 3). Although several studies have proven that a gender difference in language exists, it has to be mentioned that those studies mainly focussed on native speakers and not language learners (cf. Milroy/ Milroy 1978; Trudgill 1972).

Regarding normalisation, it is known that female speakers have an increase in F1 with a degree of vowel openness (cf. Simpson 2002). In addition, females have a smaller articulatory space, on average, which increases the speed at which they reach their target sound. However, this explanation is not applicable in this context due to the normalisation that was performed to be able to compare male and female speakers.

4.4 Limitations

This study encompasses several limitations that warrant discussion. Firstly, with regards to the normalisation tool employed, the Bark Difference Metric was utilised to establish a normalised dataset. This method, being vowel-intrinsic, effectively filters out physiological differences while preserving sociolinguistic variations. However, it is crucial to acknowledge that the normalisation process is contingent upon F3, which is utilised in the formula for calculating the normalised F1 and F2 data. In phonetics, however, F3 is affected by the position of the lips, which plays a crucial role as lip rounding is used to contrast vowels (cf. Ladefoged/Disnery Ferrari 2012). The transfer of lip rounding patterns from German to English may have influenced the participants' realisations of STRUT vowels. Additionally, words containing rhotacised vowels such as *strut*, *rushed*, or *running* had to be extracted from the dataset due to their lower F3 values (cf. Thomas/Tyler 2007).

Another factor to consider is the influence of input. The study was conducted in the canton of Bern, Switzerland, where course materials emphasise Standard British English pronunciation and spelling. However, the instructors themselves have acquired their English proficiency from various English-speaking regions, including Scotland, England, and the United States. This diversity underscores the wide range of sound production variations. Nonetheless, by employing similar course materials featuring Standard British English pronunciation, it is assumed that the students receive comparable training in that particular pronunciation.

Furthermore, it is important to note that due to the pandemic circumstances, it was not feasible for all students to utilise the same recording device. It is worth acknowledging that variations in devices may introduce potential inaccuracies in audio recordings, as demonstrated by Freeman/De Decker (2021), particularly concerning low-back vowels.

5 Conclusion

The present study was designed to determine the realisation of the STRUT vowel by Swiss German speakers in relation to the reference vowels produced in coursebooks. By focusing on two different levels of education (KV and Gymnasium) as well as male and female speakers, the findings revealed that KV produces the STRUT vowel closer to the reference target vowel. Moreover, the results suggest that male speakers perform better in the realisation of that vowel sound. Overall, it appears that besides these findings, non-native speakers are more variable in their pronunciation, resulting in between-speaker variation, which is also supported by previous research (cf. Cunningham 2008).

Research has found that the realisation of a vowel segment depends on the vowels' phonological environment and whether the sound is articulated in spontaneous speech utterances or single-word lists (cf. Thomas/Tyler 2007). Being limited to only read speech, this study lacks a

larger sample of vowels that can be compared to the reference vowels, so it would present further insight into how the participants' vowel realisation behaves in the more natural context of spontaneous speech.

The results of this study could be of significant value to English language teachers in terms of the necessity of teaching pronunciation in classrooms. Even though accents are a natural, inevitable and significant outcome of language variation, schools need to be able to teach certain guidelines for the students. Students need structures to help them with learning something new.

Overall, this study contributes to the broader field of second language acquisition in terms of L2 pronunciation.

These findings suggest several courses of action for more research to be done. On the one hand, a) how the realisation of the STRUT vowel behaves in comparison to spontaneous speech and thus includes the duration of said vowel and, on the other hand, b) a larger analysis including more speakers to check the tendencies we have found.

References

Aida, Yumie (1994): Examination of Horwitz & Cope's construct of foreign language anxiety: The case of students of Japanese. *The Modern Language Journal* 78/2: 155–168.

Alexious, Thomai (2016): "Gender Research in EFL Classrooms – Or: Are Girls Better Language Learners?" In: Elsner, Daniela/Lohe, Viviane (eds.): *Gender and Language Learning: Research and Practice*. Tübingen, Narr Francke Attempto: 85–94.

Beinhoff, Bettina (2013): *Perceiving identity through accent: Attitudes towards non-native speakers and their accents in English*. Bern: Lang.

Boersma, Paul/Weenink, David (1992): PRAAT (Version 6.3.08) praat.org [26.05.2025].

Clopper, Cynthia G./Pisoni, David B./de Jong, Kenneth (2005): "Acoustic Characteristics of the Vowel Systems of Six Regional Varieties of American English". *The Journal of the Acoustical Society of America* 118/3: 1661–1676. doi.org/10.1121/1.2000774.

Cunningham, Una (2008): "Acoustic variability in the production of English vowels by native and non-native speakers". In: Ewa Waniek-Klimczak (ed.): *Issues in accents of English*. Newcastle upon Tyne, Cambridge Scholars Publishing: 1–15.

Dong, Hanghang/Liu, Xiaofan/Zhou, Ziqi (2023): "Influential factors for gender differences in L2 learning". *Journal of Education, Humanities and Social Sciences* 8: 947–955. doi.org/10.54097/ehss.v8i.4385

Dörnyei, Zoltan (2009): "Motivation, Language Identities and the L2 Self: Future Research Directions". In: Dörnyei, Zoltan/Ushioda, Ema (eds.): *Motivation, Language Identity and the L2 Self* . Bristol, Multilingual Matters: 350–356.

Dörnyei, Zoltan/Clément, Richard (2001): "Motivational characteristics of learning different target languages: Results of a nationwide survey". In: Dörnyei Zoltan/Schmidt, Richard (eds.): *Motivation and second language acquisition*. Honolulu, University of Hawaii: 399–432.

Drummond, Rob (2010): *Sociolinguistic variation in a second language: The influence of local accent on the pronunciation of non-native English speakers living in Manchester*. PhD thesis, the University of Manchester. proquest.com/dissertations-theses/sociolinguistic-variation-second-language/docview/1314569311/se-2 [26.05.2025].

Durham, Mercedes (2007): *English in Switzerland: Inherent variation in a non-native speech community*. PhD Thesis, Université de Fribourg. folia.unifr.ch/unifr/documents/300243 [26.05.2025].

Ellis, Rod (2006): "Individual Differences in Second Language Learning". In: Davies, Alan/ Elder, Catherine (eds.): *The Handbook of Applied Linguistics*. Malden (MA), Blackwell: 525–551.

Engstrand, Olle (1988): "Articulatory correlates of stress and speaking rate in Swedish VCV utterances". *The Journal of the Acoustical Society of America* 83/5: 1863–1875.

Fabricius, Anne H. (2007): "Variation and change in the trap and strut vowels of RP: a real time comparison of five acoustic data sets2". *Journal of the International Phonetic Association* 37/3: 293–320.

Ferragne, Emmanuel/Pellegrino, François (2010): "Formant frequencies of vowels in 13 accents of the British Isles". *The Journal of the international Phonetic Association.* 40/1: 1–34.

Fischer-Jørgensen, Eli (1985): "Some Basic Vowel Features, their Articulatory Correlates, and their Explanatory Power in Phonology". In: Victoria Fromkin (ed): *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*. New York, Academic Press: 255–276.

Flege, James E. (1993): "Production and perception of a novel, second-language phonetic contrast2. *The Journal of the Acoustical Society of America* 93/3: 1589–1608.

Flege, James E. (1995): "Second language speech learning: Theory, findings, and problems". In: Strange, Winifred (ed.): *Speech perception and linguistic experience: Issues in cross-language research*. Timonium (MD), York Press: 233–277.

Flege, James E./Bohn, Ocke-Schwen (2021): "The Revised Speech Learning Model (SLM-r)". In: Wayland, Ratree (ed.): *Second Language Speech Learning.* Cambridge, Cambridge University Press: 3–83.

Flynn, Nicholas (2011): "Comparing vowel formant normalisation procedures". *York Papers in Linguistics* 11/2: 1–28.

Francis, Becky (2000): *Boys, Girls and Achievements: Addressing the Classroom Issues.* London (NY): Routledge.

Freeman, Valerie/De Decker, Paul (2021): "Remote sociophonetic data collection: Vowels and nasalization over video conferencing apps". *The Journal of the Acoustical Society of America* 149/2: 1211–1223.

Gass, Susan (2013): *Second Language Acquisition: An Introductory Course*. 4th ed. New York: Routledge.

Ge, Chunyu/Xiong, Yixuan/Mok, Peggy (2021): "How Reliable Are Phonetic Data Collected Remotely? Comparison of Recording Devices and Environments on Acoustic Measurements". *Interspeech* 2021: 3984–3988. doi.org/10.21437/Interspeech.2021-1122.

Graeppi, Christine/Leemann, Adrian (2019): "Between-speaker variation in English Learners' Realisation of Dental Fricatives". In: Calhoun, Sasha et al. (eds.): *Proceedings of the 19th International Congress of Phonetic Sciences*. Australasian Speech Science and Technology Association Inc., and International Phonetic Association: 974–978. doi.org/10.5167/UZH-185063.

Hansen Edwards, Jette, G./Zampini, Mary, L. (eds.) (2008): *Phonology and Second Language Acquisition.* Vol. 36. Amsterdam: Benjamins.

Hanulíková, Adriana/Weber, Andrea (2010): "Production of English interdental fricatives by Dutch, German, and English speakers". In: Dziubalska-Kolaczyk, Katarzyna/Wrembel, Magdalena/Kul, Malgorzata (eds.): *Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech.* New Sounds 2010. Poznan, Adam Mickiewicz Universtiy: 173–178.

Heyde, Adelaide (1979): *The relationship between self-esteem and the oral production of a second language.* PhD thesis, the University of Michigan. proquest.com/dissertations-theses/relationship-between-self-esteem-oral-production/docview/302941076/se-2 26.05.2025].

Jenner, Bryan (1989): "Teaching Pronunciation: The common core". *Speak Out!* 4: 2–4.

Kissau, Scott (2006): "Gender Differences in Motivation to Learn French". *The Canadian Modern Language Review* 62/3: 401–422.

Koul, Ravinder et al. (2009): "Multiple goal orientations and foreign language anxiety". *System* 37/4: 676–688.

Ladefoged, Peter/Johnson, Keith (2011): *A Course in Phonetics.* 6th ed. Boston (MA): Wadsworth Cengage Learning.

Ladefoged, Peter/Disnery Ferrari, Sandra (2012): *Vowels and Consonants*. 3rd ed. Chichester: Wiley-Blackwell.

Leemann, Adrian (2008): "Vowel Quality of Swiss EFL Speakers". In: Waniek-Klimczak, Ewa (ed.): *Issues in Accents of English* . Cambridge, Cambridge Scholars Publishing: 16–32.

López Rúa, Paula (2006): "The sex variable in foreign language learning: An integrative approach". *Porta Linguarum* 6: 99–114.

Milroy, James/Milroy, Lesley (1978): "Belfast: Change and variation in an urban vernacular2. In: Trudgill, Peter/Mitchel, James C. (eds.): *Social networks in urban situations*. Manchester, Manchester University Press: 19–36.

Milroy, James/Milroy, Lesley (1991): *Authority in Language: Investigating Language Prescription and Standardisation*. 2nd ed. London: Routledge.

Milroy, James/Milroy, Lesley (2012): *Authority in Language: Investigating Standard English* 4th ed. London/New York: Routledge Taylor & Francis Group.

Mori, Setsuko/Gobel, Peter (2006): "Motivation and gender in the Japanese EFL classroom". *System* 34/2: 194–210.

Mruk, Christopher J. (2013): *Self-Esteem and Positive Psychology: Research, Theory and Practice* (4th ed.). New York: Springer Publishing Company.

Munro, Murray J./Derwing, Tracey M. (1999): "Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners". *Language Learning* 49: 285–310. doi.org/10.1111/0023-8333.49.s1.8

Oxford, Rebecca (2018): "Language Learning Strategies". In: Burns, Anne/Richards, Jack C.: *The Cambridge Guide to Learning English as a Second Language*. Cambridge, Cambridge University Press: 81–90.

Peacock, Matthew/Ho, Belinda (2003): "Student Language Learning Strategies across Eight Disciplines". *International Journal of Applied Linguistics* 13: 179–200. doi.org/10.1111/1473-4192.00043

Pickering, Lucy (2006): "Current Research on Intelligibility in English as a Lingua Franca". *Annual Review of Applied Linguistics* 26: 219–233.

Piller, Ingrid/Pavlenko, Aneta (2001): "Introduction: Multilingualism, second language learning, and gender". In: Pavlenko, Aneta et al. (eds.): *Multilingualism, Second Language Learning, and Gender*. Berlin/New York, Mouton de Gruyter: 1–13.

R Core Team (2022): *A language and environment for statistical computing. R Foundation for Statistical Computing*. R-project.org/ [26.05.2025].

Rosenberger, Lukas (2009): *The Swiss English Hypothesis*. Tübingen: Francke.

Sanker, Chelsea et al. (2021): "(Don't) try this at home! The effects of recording devices and software on phonetic analysis: Supplementary material". *Language* 97/4: 360–382. doi.org/10.1353/lan.2021.0079.

Simpson, Adrian P. (2002): "Gender-specific articulatory–acoustic relations in vowel sequences". *Journal of Phonetics* 30/3: 417–435.

Spielberger, Charles D. (1983): *Manual for the state-trait anxiety inventory*. Palo Alto (CA): Consulting Psychological Press.

Strange, Winifred (2007): "Cross-language phonetic similarity of vowels: Theoretical and methodological issues2. In: Bohn, Ocke-Schwen/Munro, Murray J. (eds.): *Language Learning & Language Teaching*. Vol. 17. Amsterdam, Benjamins: 35–55.

Swan, Michael (2001): "German Speakers". In: Smith, Bernhard/Swan, Michael (eds.): *Learner English: A teacher's guide to interference and other problems*. Cambridge, Cambridge University Press: 37–51.

Thomas, Erik R. (2011): *Sociophonetics: An Introduction*. Basingstoke, Palgrave Macmillan.

Thomas, Erik R./Kendall, Tyler (2007): NORM [Computer software]. lingtools.uoregon.edu/norm/about_norm1.php [26.05.2025].

Trudgill, Peter (1972): "Sex, covert prestige and linguistic change in the urban British English of Norwich". *Language in Society* 1/2: 179–195. doi.org/10.1017/S0047404500000488.

van Son, Rob J. J. H./Pols, Louis C. W. (1990): "Formant frequencies of Dutch vowels in a text, read at normal and fast rate". *The Journal of the Acoustical Society of America* 88/4: 1683–

Walker, Robin (2022): *Teaching the Pronunciation of English as a Lingua Franca*. Oxford, Oxford University Press.

Wells, John C. (1982): *Accents of English*. 1st ed. Cambridge, Cambridge University Press.

Wester, Femke/Gilbers, Dicky/Lowie, Wander (2007): "Substitution of dental fricatives in English by Dutch L2 speakers". *Language Sciences* 29/2–3: 477–491.

Wucherer, Barbara V./Reiterer, Susanne M. (2018): "Language is a girlie thing, isn't it? A psycholinguistic exploration of the L2 gender gap". *International Journal of Bilingual Education and Bilingualism* 21/1: 118–134. doi.org/10.1080/13670050.2016.1142499.

Zarrinabadi, Nourollah/Pawlak, Mirosław (eds.) (2021): *New Perspectives on Willingness to Communicate in a Second Language*. Cham, Springer International Publishing.

Zhao, Sherry Y. (2010): "Stop-like modification of the dental fricative /ð/: An acoustic analysis". *The Journal of the Acoustical Society of America* 4/128: 1001–1024.