‘Chend’ met <e> – ‘Kind’ mit <e>: using Big Data to explore phoneme-to-grapheme mapping in Lucerne Swiss German

Urban Zihlmann[1], Adrian Leemann[1]

[1]Phonetics Lab., Department of Theoretical and Applied Linguistics, University of Cambridge
{ubz20|al764}@cam.ac.uk

Abstract

Speakers from the canton of Lucerne are infamous for spelling Middle High German (MHG) <i> as <e> when communicating in written Swiss German, e.g. *Kind* (‘child’) as <Chend>. This phenomenon has been examined only impressionistically by phoneticians. This study provides a first account of this peculiarity of Lucerne Swiss German spellers: an analysis of normalised formant frequencies of two underlyingly MHG <i> vowels from 200+ speakers of the *Dialäkt Äpp* corpus revealed that the Lucerne allophone is in reality [e] for most of the localities examined, which may explain why in vernacular writing, spellers prefer <e> over <i>. Homophony due to this peculiarity can cause misunderstandings in written and oral communication, and possibly has repercussions on the reading and writing development of Lucerne students.

Index Terms: dialectology, formants, regional variation, crowdsourcing, Swiss German, iOS, Lucerne German

1. Introduction

The canton of Lucerne (LU) has a total surface area of 1,494 km^2, and approximately 394,600 inhabitants, which makes it the biggest and most populated canton of Central Switzerland [1]. Around 86.3% of its inhabitants view German (StG) as their first language [2]. Within the SwG dialect continuum, Lucerne German is a transition zone in the centre of Switzerland [3], located between the eastern and the western dialect areas. Furthermore, LU is split by the Brünig-Napf-Reuss line (applying equally to the Aargau), which is not only regarded as a cultural border between the east and the west of German-speaking Switzerland, but also a linguistic one [4].

The most seminal work on LU SwG was conducted by [3], who provided the first grammar that included general chapters on the phonetics of the dialect. To date, however, there has been no research on one of the most salient features of LU SwG: the orthographic representation of Middle High German (MHG) <i> as <e>. To illustrate this, Figure 1 depicts a text message written by a typical LU SwG speaker:

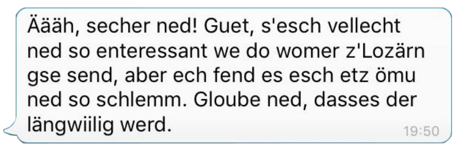

Figure1: *Text message written by a LU SwG speaker with high frequency of MHG <i> as <e>.*

The phrase reads *Ah, sicher nicht! Gut, es ist vielleicht nicht so interessant wie damals, als wir in Luzern waren, aber ich finde es ist jetzt aber nicht so schlimm. [Ich] glaube nicht, dass es dir langweilig wird*; ‘Ah, definitely not! Well, it may not be as interesting as it was when we were in Lucerne together then, but I don’t think it’s that bad now. [I] don’t think that you will be bored’. The vernacular representation features numerous MHG <i> as <e>, such as in *secher* (‘definitely’), *ned* (‘not’) etc. A vast majority of other SwG vernacular writings would spell such words as *sicher* or *nid, i.e. with <i>*. This raises the question as to why most LU speakers opt for the grapheme <e> rather than <i>. What compounds the problem at hand is that some LU speakers have been shown to represent it as <i> as well (e.g. [5, 6]).

The present study contributes to fill this gap by performing an acoustic analysis of the vowels in *Ch**e**nd/Ch**i**nd* and *tr**e**nke/tr**i**nke* (‘child’ *Kind* and ‘to drink’ *trinken*, which go back to MHG *kint* and *trinken)*. It is assumed that the MHG short vowel <i> lowered its allophones to [i], [ɪ], or [e] [3, 7]. With these analyses we try to establish whether there is an acoustic basis for LU SwG writers of the vernacular to prefer <e> rather than <i> in representing MHG <i>: we predict that for most speakers, MHG <i> is indeed realized as [e] and that for this reason, LU SwG speakers tend to map MHG <i> with <e> in writing. To test this prediction, we analysed speech data from 200+ speakers stemming from the *Dialäkt Äpp (DÄ)* corpus. As the height of a vowel strongly correlates with the first formant [8], we will primarily focus on the description of *f1*.

2. Data and methods

2.1. iOS application: ‘Dialäkt Äpp’

Dialäkt Äpp [9] enables users (1) to record 16 words and a short passage in their dialect and (2) to localise their dialect by choosing how they pronounce the 16 words in their SwG dialect. For the purpose of this study, we used functionality (1), introduced below. Prior to recording, the users of the app must indicate their age, sex, and dialect (see Figure 2, left panel).

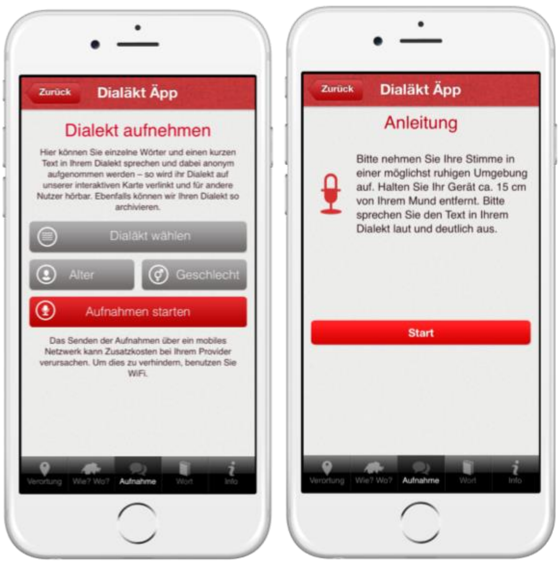

Figure 2: *User interface for dialect, age, and sex selection (left) and recording instructions (right)*

They are given instructions regarding the recording process (see Figure 2, right panel), stating: ‘Please record your voice in as quiet an environment as possible. Keep an approximate distance of about 15 cm between your device and your lips. Please articulate the text loudly and clearly in your own dialectal pronunciation’. They then record the 16 words shown on individual prompts (see Figure 3, left panel). Each iOS device from the first generation onwards has sampling rates of up to 48 kHz [10]. For the purposes of this study, 48 kHz are sufficient for reliable formant measurements, as is a sampling rate of 10 kHz [11]. After the recording process the raw wav files are uploaded on a server and tagged with unique IDs. The recordings then appear on an interactive map (Figure 3, right panel, green and purple pins). After releasing *DÄ* on 22 March, 2013, it was the most downloaded free app for iPhones [12]. Presently, it has >58,000 downloads, and its database includes c. 3,000 speakers from 452 localities across German-speaking Switzerland [13, 14].

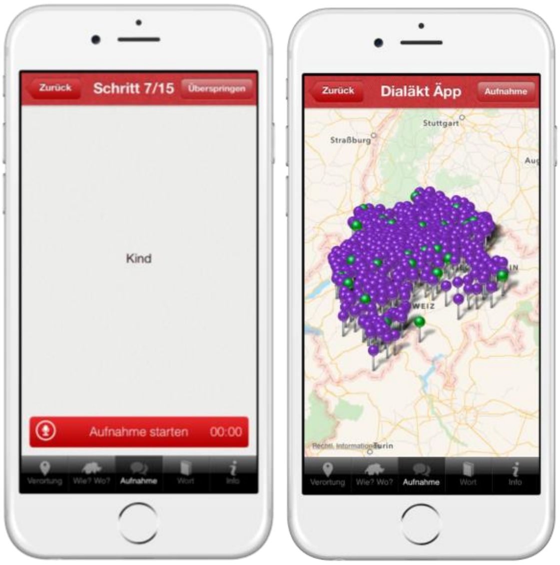

Figure 3: *User prompt for word recording (left) and interactive map of users recordings (right)*

2.2. Subjects

Users who indicated a Lucerne locality to best correspond to their dialect served as subjects. 206 speakers recorded the word *Kind* and 210 *trinken*. Speakers ranged between 10 and 77 years of age (mean=30.1; median=26.5; SD=15.0), with 47.8% males and 52.2% females. Subjects originated from virtually every corner of the canton (32 localities in total), which we divided into six regions for subsequent analyses of diatopic distributions (cf. 3.1.): Entlebuch (EB), Hinterland (HL), Lucerne-Hochdorf (L-H), Midland (ML), Mount Rigi (RG), and Schongau (SCH). The division is based on Fischer’s linguistic observations on the morphological, lexical, and phonological level [3]. For instance, EB and RG speakers show differences in vowel quantity; they articulate open-syllables such as the first syllable in *jagen* (‘to hunt’) as [ˈjɑ.g̊ə], while the rest of the canton produces them with long vowels, i.e. [ˈjɑː.g̊ə], see Figure 4.

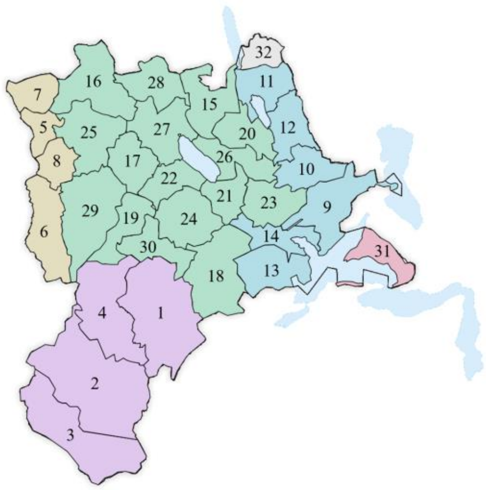

Figure 4: *Localities and broader dialect regions as used in the current sample*

2.3. Material

We chose two *DÄ* tokens with underlying MHG <i>: *Kind* ‘child’, and *trinken* ‘to drink’. Some recordings were discarded due to background noise interference or other recording errors. The percentage of discarded tokens amounted to 17.5%.

2.4. Procedure

f1 and *f2* frequencies were measured in *Praat* [15]: if the segment was >10ms, measurements were taken 10ms after the beginning of the segment (M1), 10ms before the end of the segment (M2), and in the middle of the segment (M3; see Figure 5, top panel). If the segment was <10ms, measurements were taken at the beginning (M1) and at the end (M2) of the segments, as well as in the middle (M3; see Figure 5, bottom panel). As it is unclear which temporal value is most critical in the perception of the vowels, the mean value of M1-M3 was used for the analysis.

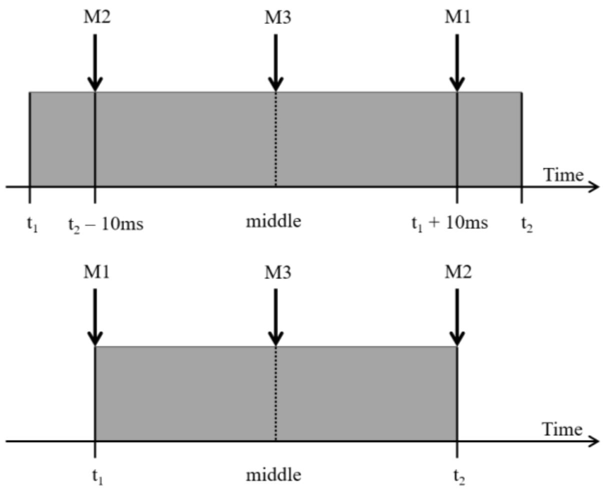

Figure 5: *Schematic of formant frequency measurements (M) >10ms (top) <10ms (bottom) (t_1 = beginning of the segment; t_2 = end of the segment)*

We normalised formant measurements using Bladon et al.’s base formula [16] which, however, only accounts for differences in adult males and females. Thus, we adapted the formula to enable comparisons with younger speakers. To this end we considered the estimated vocal tract lengths of men and women (based on [17]) and calculated the age-appropriate amount of Barks to be subtracted from Bladon et al.'s formula. The difference between the average vocal tract length of an adult male and an adult female is 28.4 mm (m=169.3 mm; f=140.9 mm) and the difference between the respective value subtracted from Bladon et al.'s formula is 1.0 Bark (-0.53 Bark for the males; -1.53 Bark for the females). This allows us to calculate the millimetre-to-Bark ratio per millimetre difference to the mean adult vocal tract length, which is 0.035 Barks, i.e. $\frac{1}{28.4}$. We then included this as a subtraction term in Bladon et al.'s equation. This results in formula (1) for male and (2) for female speakers. The variables to be filled in are the raw formant frequencies in Hertz (f_i) and the mean vocal tract length by age (VTL_{age}).

$$(1)\quad f_i^N = 26.81\left(\frac{f_i}{1960+f_i}\right) - 0.53 - \left(\frac{1}{28.4}\left(169.3 - [VTL_{age}]\right)\right)$$

$$(2)\quad f_i^N = 26.81\left(\frac{f_i}{1960+f_i}\right) - 1.53 - \left(\frac{1}{28.4}\left(140.9 - [VTL_{age}]\right)\right)$$

As the equation results in Barks scores, we retransformed it to Hertz with *hqmisc* [18] (which uses Traunmüller's [19] formula) since the *R* package for plotting the vowels (*phonR* [20]) operates on the Hertz scale. Statistical analyses were conducted using *RStudio* [21].

3. Results

3.1. Diatopic differences

Table 1 summarises the mean formant frequencies and standard deviations (SD) by location.

	Locality	Mean *f1*	Mean *f2*	SD *f1*	SD *f2*	Area
1	Entlebuch	355	2069	42.6	192.6	EB
2	Escholzmatt	297	1934	69.4	65.4	EB
3	Marbach	355	1787	6.8	310.3	EB
4	Schüpfheim	375	2020	46.9	167.1	EB
5	Altbüron	427	2014	46.8	135.3	HL
6	Luthern	456	1823	10.1	87.8	HL
7	Pfaffnau	371	2015	9.5	38.3	HL
8	Zell	370	2112	58.1	186.6	HL
9	Ebikon	360	2018	41.8	183.3	L-H
10	Eschenbach	353	2001	89.3	200.4	L-H
11	Hitzkirch	381	2090	69.8	177.3	L-H
12	Hohenrain	432	2106	66.7	80.5	L-H
13	Horw	390	1902	43.4	203.6	L-H
14	Luzern	365	2021	50.9	188.4	L-H
15	Beromünster	367	2012	30.1	176.5	ML
16	Dagmersellen	386	1942	86.0	154.4	ML
17	Grosswangen	334	2087	27.3	185.1	ML
18	Malters	399	2132	58.4	123.8	ML
19	Menznau	442	2114	43.4	101.9	ML
20	Neudorf	358	1792	53.1	304.1	ML
21	Neuenkirch	378	2109	22.3	165.8	ML
22	Nottwil	347	1989	28.6	101.3	ML
23	Rothenburg	379	2019	58.5	186.0	ML
24	Ruswil	363	1951	44.1	161.5	ML
25	Schötz	373	2065	58.5	108.5	ML
26	Sempach	382	2140	34.0	151.7	ML
27	Sursee	365	1911	55.6	273.7	ML
28	Triengen	360	2057	51.4	139.3	ML
29	Willisau	402	2088	66.3	138.2	ML
30	Wolhusen	369	2071	44.4	155.6	ML
31	Weggis	353	1947	53.0	130.6	RG
	Total	376	2011	47.3	160.5	

Table 1: *Normalised vowel frequencies of MHG <i> and SDs by locality*

Overall, the mean *f1* frequency for the entire canton of LU is 376 Hz (SD=47.3 Hz). The lowest *f1*s (i.e. the highest articulations) are found in Escholzmatt (297 Hz), followed by Grosswangen (334 Hz), Nottwil (347 Hz), Weggis, and Eschenbach (both 353 Hz). The highest *f1*s (i.e. the lowest articulations) were found in Luthern (456 Hz), Menznau (442 Hz), Hohenrain (432 Hz), Altbüron (427 Hz), and Willisau (402 Hz). Vowel height seems to be rather stable throughout the canton (SD=47.3 Hz).

3.2. Differences by area

Table 2 summarises the mean formant frequencies and SDs by area; Figure 5 shows the values on the f1 / f2 vowel pane.

Area	Mean *f1*	Mean *f2*	SD *f1*	SD *f2*
EB	344	2004	54.8	184.7
HL	397	2043	57.3	170.9
L-H	367	2016	55.1	189.3
ML	376	2024	56.6	185.8
RG	340	1948	39.3	141.1

Table 2: *Mean normalised vowel frequencies of MHG <i> (in Hz) by area*

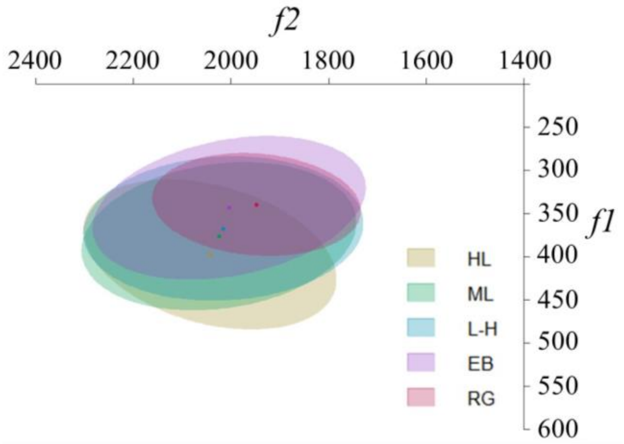

Figure 6: *Vowel ellipses of mean f1 and f2 frequencies with the corresponding SD (diameter of the oval)*

Figure 6 reveals substantial overlap between the regions. On the *f1* pane, RG reveals the lowest SD (39.3 Hz), while in HL, we observe most variation in *f1* (57.3 Hz). The highest articulation of MHG <i> is found in the RG area (340 Hz), whereas the lowest variant is found in HL (397 Hz). Both ML and L-H are in the vicinity of HL's values (ML, 376 Hz; 21 Hz lower than HL; L-H, 367 Hz; 30Hz difference to HL). EB, too approaches these values (344 Hz), although they produce a higher variant. Taken together, the northern three areas in the cantons all lie within a range of 30 Hz for *f1*, which accounts for the overlap in Figure 6. All areas exhibit values that approximate mean frequencies of [e] of 390 Hz as suggested by Catford (as opposed to 240 Hz for [i]), but the linguistic background of the male speaker remains unspecified [22]. When data from StG are considered, such as Reubold [23], who found the formant frequencies of [e] to be 299 Hz, and 259 Hz for [i], the articulation in the entirety of LU seems to take place even lower.

4. Discussion

Our findings suggest that – on the whole – LU SwG articulations of MHG <i> are closer to [e] rather than [i]. There are regional differences, however: RG and EB demonstrate the highest variants, which has been previously documented in [3]. In Grosswangen and Nottwil, both within ML, however, we also found high articulations – yet their production is slightly lower than in RG and EB. Generally, however, the most suitable allophonic representation for MHG <i> appears to be [e]: here, mean *f1* frequencies are all in the vicinity of Catford's values for [e], and even higher (i.e. LU SwG articulates MHG <i> even lower) than the ones suggested by Reubold.

There are a number of implications to these findings. This lowering can cause confusion when LU speakers write to non-LU speakers in SwG vernacular, such as in informal texting or emails (see Figure 1). The formant frequencies reported in this study suggest that LU speakers tend to produce MHG <i> as [e], albeit with between-locality variation. If the writer chooses to represent this allophone with the grapheme <e>, misunderstandings could occur. If, for example, Zurich (ZH) SwG speakers read the message shown in Figure 1, they would likely associate <e> with the phonemes /ɛ/, /eː/, and /ə/, rather than conceiving of them as variants of MHG <i>, as intended by the LU SwG writer. Aside from potential confusion in written communication, in verbal communication, too, new homophones may emerge due to the lower articulation in LU SwG: the words *mer* ('me' *mir*), *mer* ('we' *wir*), and *Meer* ('sea' *Meer*) can all be homophonous and articulated as [meːɾ] in LU SwG. Moreover, LU SwG equivalents for the words 'seen' *gesehen* and 'been' *gewesen* are both neutralised to [g̥s̥eː], while ZH speakers maintain the [g̥s̥eː] / [g̥s̥iː] contrast. Though in isolation these words may cause misunderstandings, phrasal context typically resolves this.

The fact that the majority of LU dialect speakers use [e] for MHG <i> could also have implications for the classroom setting. German-speaking Switzerland is diglossic, yet LU children typically do not receive formal StG education until they begin school or kindergarten at age 5. By then, they will have learned to speak SwG vernacular, but will not have mastered the orthography of StG. As they grow older, they will first spell words close to what they sound like [24], followed by a simple grapheme-phoneme correspondence mechanism that will start to emerge at around age 7 [25]. However, when a given grapheme has more than one corresponding sound, or in other words, when the phoneme-grapheme correspondence is not 1:1, the spelling and reading acquisition process may be decelerated to some degree. This has been reported for English and Turkish students. When a student's native language has an irregular phoneme-grapheme correspondence as in English, they will typically master reading and spelling later than students whose native language has a more reliable sound-to-letter correspondence, such as in Turkish [26]. In the context of SwG, LU students will have to become aware that some of the [e]s they produce in SwG are orthographically represented by <e>, and some by <i> in StG – albeit vernacular writing allows for many (idiosyncratic) degrees of freedom. A speaker of ZH SwG, for example, who appears to have a more straightforward mapping of [i] to MHG <i> does not encounter this issue.

Interestingly, SwG speakers from western German-speaking Switzerland feature lowered MHG <i> as well, e.g. Bern (BE) German [27, 28]. Yet, they typically use <i> in written vernacular writing (e.g. <Chind> for *Kind,* 'child'). This suggests that LU SwG speakers conceptualise MHG <i> differently from these speakers, using an alternate strategy for phoneme to grapheme mapping. Further research is needed to explore (a) whether BE SwG speakers, in reality, have equally low articulations of MHG <i> as LU SwG speakers do and (b) whether BE and LU SwG perceive vowels equally. An exploration of both of these issues would help us better understand the peculiarity of LU SwG speakers' phoneme-to-grapheme mapping.

5. Conclusion

The findings of this study suggest that for most LU SwG dialects, the production of MHG <i> is closer to [e] rather than to [i]. Results on a more regional level revealed that speakers in the northern parts of the canton tend to articulate the phoneme closer to [e], while *f1* frequencies of RG and EB suggest the allophone to be somewhat higher for these regions (as reported in [3]). We speculate that misunderstandings may arise due to this dialect-specific phoneme-to-grapheme mapping when LU speakers are in written contact with non-LU speakers, e.g. in informal text messages. This lowering may have implications on the spelling acquisition process of StG in LU primary school students, given that students have to learn to dissociate LU-specific [e] from MHG <i>.

6. Acknowledgments

This study is one part of Zihlmann's MPhil thesis at the University of Cambridge. He wishes to thank the *Geert-und-Lore-Blanken-Schlemper Foundation* in CH-6210 Sursee, the *Hans-und-Wilma-Stutz Foundation* in CH-9102 Herisau, *Prof Beat Siebenhaar* (Leipzig), *Prof Francis Nolan (Cambridge)*, and *Marie-José Kolly (Zurich)* for their support and for co-creating the *DÄ* corpus.

7. References

[1] LUSTAT Statistics Lucerne. (2015). Bevölkerung nach Kanton. Retrieved 01 May 2016 from https://www.lustat.ch/files_ftp/daten/kt/0003/w011_001t_kt0003_zz_d_0000.html.

[2] Federal Statistics Office. (2014). *Languages and religions Data*. Retrieved 30 April 2016 from http://www.bfs.admin.ch/bfs/portal/en/index/themen/01/05/blank/key/sprachen.html.

[3] Fischer, L. (1960). Luzerndeutsche Grammatik und Wegweiser zur guten Mundart. Zurich: Schweizer Spiegel Verlag.

[4] Weiss, R. (1947). Die Brünig-Napf-Reuss-Linie als Kulturgrenze zwischen Ost- und Westschweiz auf volkskundlichen Karten. Geographica Helvetica 2(3), 153-75. doi:10.5194/gh-2-153-1947.

[5] Hotzenköcherle, R., Bigler, N., Schläpfer, R., & Börlin, R. (1984). Die Sprachlandschaften der deutschen Schweiz. Aarau: Sauerländer.

[6] Zihlmann, J. (1975). De jung Chuenz und anderi Gschichte. Luzern: Murbacher-Verlag.

[7] Haas, W. (1978). Sprachwandel und Sprachgeographie: Untersuchung zur Struktur des Dialektverschiedenheit am Beispiele der schweizerdeutschen Vokalsysteme. Wiesbaden: Steiner.

[8] Ashby, M., & Maidment, J. A. (2005). Introducing phonetic science. Cambridge: Cambridge University Press.

[9] Leemann, A., and Kolly, M.-J., Dialäkt Äpp. https://itunes.apple.com/ch/app/dialakt-app/id606559705?mt=8, 2013.
[10] iPhone Specifications (2016). Retrieved 17 June 2016 from http://www.everyiphone.com/.
[11] Podesva, R. J., & Zsiga, E. (2013). Research methods in linguistics. Cambridge, UK: Cambridge University Press, 169-194.
[12] http://www.appannie.com/.
[13] Leemann, A., Kolly, M.-J., Purves, R., Britain, D., & Glaser, E. (2016). Crowdsourcing Language Change with Smartphone Applications. PLoS ONE, 11(1): e0143060. doi:10.1371/journal.pone.0143060.
[14] Leemann, A. (under review). Analyzing dialectal variation in articulation rate using crowdsourced speech data. Journal of Linguistic Geography.
[15] Boersma, P. & Weenink, D. (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.17, retrieved 21 April 2016 from http://www.praat.org/.
[16] Bladon, R. A. W., Henton, C. G., & Pickering, J. B. (1984). Towards an auditory theory of speaker normalization. Language and Communication, 4(1), 59-69.
[17] Goldstein, U. G. (1980). An articulatory model for the vocal tracts of growing children. PhD dissertation, Massachusetts Institute of Technology, Cambridge, MA.
[18] Quené, H. (2014). *hqmisc* - Miscellaneous convenience functions and dataset. Version 0.1-1.
[19] Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. Journal of the Acoustical Society of America, 88(1), 97-100.
[20] McCloy, D. R. (2015). *phonR* - Tools for Phoneticians and Phonologists. Version 1.0-3.
[21] RStudio Team (2015). RStudio: Integrated Development for R [Computer program]. RStudio, Inc., Boston, MA. Version 0.99.486, retrieved 19 October 2015 from http://www.rstudio.com/.
[22] Catford, J. (2001). A practical introduction to phonetics. Oxford: Clarendon.
[23] Reubold, U. (2012). Über die Zusammenhänge zwischen Grundfrequenz und Vokalhöhe: Evidenzen aus longitudinalen Altersstimmenstudien, Perturbations- und Vokalerkennungsexperimenten. München.
[24] Ehri, L. (1987). Learning to read and spell words. Journal of Literacy Research HJLR, 19(1), 5-31. doi:10.1080/10862968709547585.
[25] Marsh, C., Friedman, M., Welch, V., & Desberg, P. (1981). A cognitive-developmental theory of reading acquisition. In G. E. MacKinnon & T. G. Waller (eds.), Reading research: Advances in theory and practice, Vol. 3. New York: Academic.
[26] Oney, B., & Goldman, S. R. (1984). Decoding and comprehension skills in Turkish and English: Effects of the regularity of grapheme-phoneme correspondences. Journal of Educational Psychology, 76(4), 557-568. doi:10.1037/0022-0663.76.4.557.
[27] Leemann, A., M.-J. Kolly, F. Nolan (2015). It's not phonetic aesthetics that drives dialect preference: the case of Swiss German. Proceedings of ICPhS 2015, 08/2015 Glasgow (UK).
[28] Marti, W. (1985). Berndeutsch-Grammatik für die heutige Mundart zwischen Thun und Jura. Bern: Francke.