DOI: 10.17469/O2108AISV000017

ELISA PELLEGRINO

Dynamics of short-term cross-dialectal accommodation. A study on Grison and Zurich German

This study investigates whether rhythmic features are object of accommodation between Grison and Zurich German (henceforth GRG and ZHG) speakers, insomuch as it was previously observed for vowel formants. Cross-dialectal rhythmic accommodation and its evoking/inhibiting factors (e.g., acoustic distance vs dialect markedness, new vs previously heard words) were examined in a corpus of pre- and post-dialogue recordings, performed by 18 pairs of GRG and ZHG speakers. Three rhythmic measures were designed which were based on cross-dialectal timing differences related to intervocalic sonorants gemination, open syllable lengthening and reduction of word-final vowels. Rhythmic accommodation was quantified measuring the acoustic distance in the realization of the three durational contrasts before and after dialogical interactions. Results revealed that, unlike vowel formants, rhythmic features did not evoke cross-dialectal adjustments under the given experimental circumstances.

Keywords: Phonetic Convergence, Speech Rhythm, Dialects in Contact, Speech Production.

1. *Introduction*

Our daily experience with vocal signals tells us that the way individuals sound like they do is extraordinarily variable. Individuals can unintentionally change their voice as a result of emotional states or fatigue, but can also adapt their vocal repertoire to respond to background conditions (e.g., noise vs quiet, Summers, Pisoni, Bernacki, Pedlow & Stokes, 1988; Traunmüller, Eriksson, 2000), interlocutors' age and linguistic origin (e.g., Kemper, Finter-Urczyk, Ferrell, Harden & Billington, 1998; Kuhl, Andruski, Chistovich, Chistovich, Kozhevnikova, Ryskina, Stolyarova, Sundberg & Lacerda, Ferguson, 1997; Ferguson, 1975), or to convey personality traits (e.g, Hughes, Mogilski & Harrison, 2014; Carli, LaFleur & Loeber, 1995).

A phenomenon in speech communication which also affects the way we speak is vocal accommodation (Giles, 2008). This phenomenon is alternatively known as alignment (Pickering, Garrod, 2006), entrainment (Brennan, 1996), synchrony (Edlund, Heldner & Hirschberg, 2009), mimicry (Pentland, 2008) and chameleon effect (Chartrand, Bargh, 1999) and describes the tendency of interlocutors to adjust their verbal and non-verbal behaviour during interactions or after exposure to a communication partner or a model talker (e.g, Nguyen, Delvaux, 2015; Pardo, 2012; Ruch, Zürcher & Burkart, 2017 for review).

Two major theoretical frameworks have been proposed to explain accommodation. One is the Communication Accommodation Theory (CAT) (e.g., Giles, Mulac,

Bradac & Johnson, 1987), which holds that speakers accommodate to their interlocutors, either by becoming more similar (convergence) or by accentuating individual differences (divergence). In this view, convergence and divergence signal speakers' intention to express social closeness to or distance from their own interlocutors. In contrast to the social-oriented account of accommodation, there is the Interactive Alignment Model (IAM) proposed by Pickering & Garrod (2004). They argue that convergence is primed by an automatic, bidirectional link between speech perception and production that enhances mutual understanding between interlocutors. Evidence in support of both models exists. Recently, a hybrid approach has been proposed that sees accommodation as an unconscious phenomenon mediated by a range of social and linguistic constraints (e.g., Babel, 2012; Pardo, 2012; Pardo, Urmanche, Wilman & Wiener, 2017).

An example of accommodation which is linguistically and socially selective is the one observed between speakers of different dialects. Studies on cross-dialectal accommodation have, indeed, shown that more convergence is evoked when:

– the distance between the phonetic repertoires of study participants is large (e.g., Babel, 2010, 2012; Walker, Campbell-Kibler, 2015; Ross, Lilley, Clopper, Pardo & Levi, 2021);
– cross-dialectal variants are perceptually salient (e.g, Mitterer, Muesseler, 2013; MacLeod, 2012; Troncoso-Ruiz, Elordieta, 2017), but less socially stereotyped (e.g., Babel, 2010; Walker, Campbell-Kibler, 2015; Clopper, Dossey, 2020);
– speakers have positive attitudes toward the residents of a dialectal region (Babel, 2010).

Acoustically, cross dialectal convergence has been typically measured in terms of vowel formants (e.g. Babel, 2010, Walker, Campbell-Kibler, 2015; Clopper, Dossey, 2020; Ross et al., 2021), word duration (Clopper, Dossey, 2020), word initial or final consonants/consonant clusters (e.g. MacLeod, 2012; Mitterer, Muesseler, 2013, Ross et al., 2021), and intonation patterns (e.g. D'Imperio, Cavone & Petrone, 2014; Troncoso-Ruiz, Elordieta, 2017; Romera, Elordieta, 2013). In terms of languages under scrutiny, most studies have been conducted in regional varieties of English (see a.o. Evans, Iverson, 2007; Babel, 2010; Clopper, Dossey, 2019, 2020), but research is also available for varieties of Spanish (e.g., MacLeod, 2012; Troncoso-Ruiz, Elordieta, 2017), French (e.g., Lelong, Bailly, 2010; Bullock, Gerfen, 2004) and Italian (D'Imperio et al., 2014).

Research on accommodation has been conducted also on Swiss German dialects (Werlen, Schlegel, 2006; Ruch, 2015, 2021; Pellegrino, Schwab & Dellwo, accepted), and the present study aims at advancing the knowledge of forms and factors affecting acoustic convergence between the dialectal varieties spoken in the Cantons of Zurich and Grisons.

1.1 Previous research on accommodation between Swiss German dialects

In a longitudinal study on 18 speakers from southern Switzerland relocated to Berne, Werlen & Schlegel (2006) showed that the amount of convergence towards

the Bernese variety was largely accountable for by the speakers' attitude towards their home or host region. Those who intended to go back or had fewer positive feelings towards the host region exhibited lower degrees of convergence. Ruch (2015) examined vowel convergence and divergence between pairs of GRG and ZHG speakers in pre- and post-dialogical items extracted from a picture naming task. She observed that ZHG speakers converged more to GRG speakers than vice versa, especially in low vowels and in words which served as stimuli in dialogical interactions. The same pattern, however, did not replicate when she investigated cross-dialectal accommodation including in the analysis items elicited through a different task (story retelling) (Ruch, 2021). The lack of convergence was interpreted in the light of the sociolinguistic status of Swiss German dialects which are mutually intelligible and share equal prestige. The author also used an agent-based modelling to simulate the change in vowel space in the absence of social and interactional factors. The results of the simulation showed that convergence was dependent on the variability of vowels. The dialect with the more dispersed distribution converged towards the dialect with the more compact distribution.

It has been shown, however, that patterns of convergence in one measure within a pair or within a speaker cannot be taken to be representative of pairs and speakers' overall convergence patterns in other measures (Sanker, 2015; Cohen Priva, Sanker, 2018). Therefore, in this study we tested whether patterns of vowel convergence between GRG and ZHG speakers identified in Ruch (2015)[1] would replicate for another set of acoustic features which varies between the two dialects, namely rhythmic properties. A preliminary investigation was conducted that compared the rhythmic distance between pairs of GRG and ZHG speakers before and after dialogical interactions (Pellegrino et al., accepted)[2]. The results of this study pointed to maintenance, in other words the rhythmic distance between pairs did not change substantially after participating to dialogical interactions, nor was it influenced by the degree of dialect markedness of the examined features. In view of the hypothesized influence of short-term accommodation on language variation and change (Trudgill, 1986), from this study only, however, it would have been premature to make any conclusion regarding the contribution of rhythmic patterns and the two dialects to the diffusion of linguistic innovation and dialectal levelling

[1] The study on cross-dialectal rhythmic accommodation was designed, carried out and submitted to Studi AISV series before the paper by Ruch (2021) had been available online (10 September 2021). Therefore, the research questions and hypotheses of the present study are based on the findings reported in Ruch (2015).

[2] A detailed explanation as to the reasons why rhythmic features are supposed to be object of mutual adaptations is found in Pellegrino et al. (accepted). To put it simply, we based our assumption on evidence showing (1) rhythmic alignment in synchronous reading tasks (Cerda-Oñate, Toledo Vega & Ordin, 2021), and (2) rhythmic adjustments as function of interlocutors' age and cognitive development (for comparisons between the rhythm of speech directed to infants and to adults, see Payne, Post, Astruc, Prieto & Vanrell, 2009; Leong, Kalashnikova, Burnham & Goswami, 2017). We took these findings as evidence that speakers may also adjust their rhythmic behavior after being exposed to a dialogue partner speaking a language with a distinct rhythmic pattern.

in German speaking Switzerland. The analysis of accommodation performed in Pellegrino et al. (accepted), indeed, took the behaviour of the pair as whole. No distinction was made between items which were used during the diapix tasks and those which were not. Moreover, this type of analysis based on the pair is not particular informative in relation to the behaviour of individual ZHG and GRG speakers after being exposed to each other's dialect. In the paper by Ruch (2015), indeed, accommodation within pair was documented but remarkable variability between speakers, speaker pairs and items was also observed. To gain an in-depth comprehension of dynamics of cross-dialectal rhythmic accommodation, in the present paper we compared pre- and post-dialogical productions within individual GRG and ZHG speakers in a pair (cf. §2.2) and distinguished between new items and those which served as the basis of the dialogical interactions.

Before turning to the specifics of the present study, however, we will first:

1. sketch the sociolinguistic situation of German speaking Switzerland (cf. §1.2);
2. present the cross-dialectal phonetic differences which legitimate the assumption of inter-speaker rhythmic accommodation examined in the current research (cf. §1.3).

1.2 The sociolinguistic situation of German Speaking Switzerland

The sociolinguistic situation of German speaking Switzerland is characterized by medial diglossia. Standard German is used especially in writing, while Swiss German dialects are the default means of daily speech communication. The use of Swiss German dialects is not stigmatized in favour of standard variants. They are employed in public domains, in the media, in education (Scherrer, Samardžić & Glaser, 2019), as well as in inter-dialectal setting due to their highly mutual intelligibility.

Swiss German dialects share equal prestige (none of the dialect has the status of standard variety), whereas they rather vary in popularity, social attributes, and degree of perceived pleasantness (Ruch, 2018). For example, Bernese dialect is described as 'colorful', 'cozy', 'countryside-ish'; Thurgau dialect as 'bright', 'hard', 'unpopular' or even 'poisonous'. With respect to the two varieties examined in the present study, ZHG is perceived as 'arrogant', 'unsympathetic', 'balanced,' 'common,' 'monotonous,' and 'dislikeable', while GRG as 'original,' 'warm,' and 'beautiful' (for the attributes associated to dialects mentioned above, cf. studies in Leemann, Kolly & Nolan 2015, Ruch, 2018). It will be interesting to see whether these asymmetric attitudes towards GRG and ZHG will be reflected in the direction of cross-dialectal rhythmic accommodation.

1.3 Cross-dialectal differences between Zurich and Grison German

GRG and ZHG present noticeable phonetic differences (see a.o., Eckhardt, 1991; Fleischer, Schmid, 2006; Leemann, 2012). These have to do with:

– the vowel system, principally in the quality of front vowels;
– word-initial and post-vocalic k, realised in GRG as aspirated [kh], and in ZHG either as velar fricative [x] or velar affricate [kx];

– the speech rate that is lower in GRG than in ZHG;
– intonation contours, with GRG showing high pitch-phrase initially and more marked declination than ZHG.

The two dialects also exhibit crucial segmental durational differences that lead to a different rhythmic organisation. These differences have mainly to do with:

– intervocalic sonorants gemination in words ending in -e, (henceforth ISG);
– open syllable lengthening (henceforth OSL);
– vowel reduction in word final position (henceforth RedVow).

Given that segmental timing properties are among the acoustic correlates of speech rhythm, in this paper we will refer to the three cross-dialectal differences in ISG, OSL and RedVow as rhythmic differences. In GRG intervocalic sonorants can be realized either as geminate or as singleton consonants, while in ZHG only as singletons (e.g., 'Sun': GRG > ['sunnɐ]/[sunɐ]; ZHG > ['sunə]). In GRG, open syllables can be either lengthened or not, while in ZHG the tendency to lengthen the open syllable has not been documented (e.g., 'Sole': GRG> ['so:lɐ]/['solɐ]); ZHG > ['solə]). In GRG, vowels in word-final position do not undergo vowel reduction in quality, and presumably either in duration, while in ZHG word-final vowels are always reduced (i.e., 'Soup' GRG> ['suppɐ]; ZHG> ['suppə]). That the differences in the quality of ending vowels also accompany distinct duration patterns is supported by findings in Leemann, Dellwo, Kolly & Schmid (2012). Here it was documented that the group of Midland dialects (to which ZHG belongs to) exhibited higher durational variability of vocalic intervals than the group of Alpine dialects (to which GRG belongs to). The lower durational variability of vocalic intervals of Alpine dialects was accounted for by the tendency of this dialectal group of retaining full vowels in unstressed position.

1.4 Research questions and hypotheses

In the present study, we examined whether:

– speakers of GRG and ZHG converge in segmental timing properties in the same direction as for vowel quality;
– factors like acoustic distance/dialect markedness/word type (new or previously heard items) account for patterns of accommodation within individual GRG and ZHG speakers.

If patterns of vowel convergence identified in Ruch (2015) replicate for segmental temporal properties, we expect that ZHG speakers converge more to their dialogue partners, while GRG speakers persist in their original behaviour (maintenance). Considering findings showing that speakers converge more for features that differ mostly between dialects (e.g., MacLeod, 2012; Ruch, 2015) and between the speakers and the model talkers (e.g., Babel, 2012), we hypothesise that more accommodation is evoked by RedVow than ISG and OSL. RedVow, indeed, is one of the features that best distinguishes the two dialects. GRG indeed exhibits open syllable lengthening – though in articulatory contexts other than ZHG – and presents longer nasal duration in -CCer words. Nevertheless, the realisation

of reduced vowels is also a strong dialect marker for ZHG (Ruch, 2018). In view of evidence about less convergence for features that are dialect markers (e.g., Babel, 2010; Walker, Campbell-Kibler, 2015; Clopper, Dossey, 2020), we cannot exclude the possibility of speakers diverging or maintaining their original behaviour. Following the same line of reasoning (less convergence for dialect markers), between OSL and ISG the former should be less prone to accommodation since it is also perceived as another typical feature for GRG (Ruch, 2018). Based on findings showing more convergence for items previously heard than for new items (e.g., Goldinger, 1998; Nielsen, 2011; Ruch, 2015), we also expect more convergence in ZHG speakers for the items which were previously utilized in the diapix tasks.

2. *Method*

The speech material and the method for calculating the scores of the three rhythmic measures are identical to the study by Pellegrino et al. (accepted). The operationalization of accommodation and the examination of factors affecting convergence, however, vary between the studies. In Pellegrino et al. (accepted), we quantified accommodation at the level of the pair, viz., we compared the Euclidean distance in the three ratio measures between members of a dyad in lexical items produced before and after two dialogical interactions. We also tested the effect of rhythmic measures (ISG, OSL and RedVow) on the amount of convergence. In the current study, instead, we quantify accommodation within individual GRG and ZHG members of a pair (henceforth accommodation within a speaker, cf. §2.2), and we test the interaction and the main effect of multiple factors evoking or inhibiting cross dialectal convergence (i.e., dialect, rhythmic measures, word type).

2.1 Speech Material

The corpus utilized in this study consisted of:

- audio-recorded diapix tasks (i.e., problem-solving 'spot the difference' picture tasks, cf. Van Engen, Baese-Berk, Baker, Choi, Kim & Bradlow, 2010) performed by 18 pairs of previously unacquainted GRG and ZHG female speakers.
- 18 pre- and 18 post-dialogue recordings (picture naming task and retelling a story based on a comic), performed individually by GRG and ZHG participants.

The diapix tasks were designed to elicit the target words present in picture naming task and story retelling. The present study reports on the data extracted from the picture naming task (henceforth PNT). As explained in Pellegrino et al. (accepted), the choice of selecting data from PNT was motivated by the fact that this permitted to control for the effect of the item variability in the assessment of: (a) cross-dialectal differences before the interaction; (b) differences in distance between ZHG and GRG speakers before and after the interaction. For the list of lexical items used in this study and the dialectal features they instantiate, cf. Tab. 1.

Table 1 - *List of items from PNT. Items in bold were utilized also in the diapix tasks (Adapted from Pellegrino et al., accepted)*

ISG	OSL	RedVow
Brunnen, ***Pfanne****,* ***Sonne****, Spinne,* ***Welle***	*Besen, Esel, Graben, Käfer, Lupe,* ***Nase****,* ***Schlafen****, Melone*	*Besen, Brunnen,* ***Flosse****,* ***Graben****,* ***Lampe****, Lunge,* ***Lupe****, Melone,* ***Nase****,* ***Pfanne****,* ***Schlafen****,* ***Sonne****, Spinne,* ***Suppe****,* ***Welle***

2.2 Data Analysis and Statistics

To understand whether GRG and ZHG speakers produce the rhythmic features more similarly after participating in the diapix tasks, the following steps were taken:

– From the pre- and post-dialogue recordings of individual speakers in PNT, we extracted the lexical items instantiating the three target rhythmic features (ISG, OSL and RedVow).
– For every item, we automatically measured the duration of individual segments.

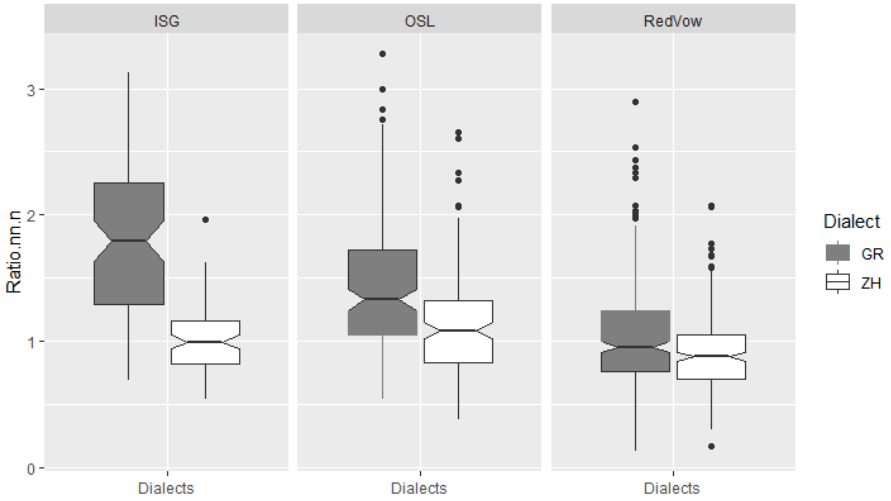
We used the raw measures of segment duration to calculate the three ratio measures which were designed *ad hoc* to capture cross-dialectal differences in ISG, OSL and RedVow[3].

– For ISG, we calculated the ratio between the duration of intervocalic sonorants (l, n) in -CCe words (e.g., *Sonne, Welle*) and that of the corresponding sonorant in -Ce words (l or n from the item *Melone*).
– For OSL, we calculated the ratio between the duration of stressed vowels in open syllables and that of the unstressed vowel within the same item.
– For RedVow, we calculated the ratio between the duration of stressed vowels in open and closed syllables and that of the unstressed vowel within the same item.

As shown in Pellegrino et al. (accepted), the speakers of the two dialects realised the three durational contrasts differently before the interaction, and this ensured that there was room for accommodation. GRG speakers scored significantly higher values as compared to ZHG speakers (Fig. 1). This was expected for ISG and OSL but not for RedVow. In GRG, indeed, intervocalic sonorants can be pronounced also as geminates, and open syllables can be lengthened, but the reduction of unstressed vowel is not that marked as in ZHG. This was, instead, imputed to the fact ZHG speakers might have not drastically reduced the duration of unstressed vowels in word-final position, as these vowels were subjected to pre-pausal lengthening.

[3] We voluntarily refrained from using the so-called rhythmic metrics for the quantification of rhythmic accommodation since these have been typically applied to corpora of utterances with a sufficiently large amount of consonantal and vocalic intervals. The dataset used in this study, instead, consisted of individual bi-syllabic lexical items with two consonantal and vocalic intervals (except for *Melone*, which has three). These were deemed to be insufficient for an accurate analysis using rhythmic metrics.

Figure 1 - *Cross dialectal differences in ISG (left) and OSL (centre), RedVow (right) in pre-dialogue PNT (Pellegrino et al., accepted)*



To quantify acoustical patterns of accommodation between GRG and ZHG speakers, we adapted methods typically utilized to quantify vowel convergence and divergence (Babel, 2012) to the rhythmic measures under study. Therefore, we measured the so-called accommodation within an individual (henceforth DD_speak) by calculating the following distance measures:

– Distance 1 (d1): Euclidean distance in the score of a ratio measure between GRG and ZHG speakers' pre-dialogue production of a lexical item. If we take '*Brunnen*' and ISG ratio as an example, d1 is calculated as follows: GRG pre $_{[ISG - Brunnen]}$ – ZHG pre $_{[ISG - Brunnen]}$.
– Distance 2 (d2): Euclidean distance in the score of a ratio measure between one speaker's post-dialogue production and her dialogue partners' pre-production of the same lexical item as in d1. So, if we consider the example above with '*Brunnen*' and ISG, d2 is calculated as follows: GRG post $_{[ISG - Brunnen]}$ – ZHG pre $_{[ISG - Brunnen]}$, and ZHG post $_{[ISG - Brunnen]}$ – GRG pre $_{[ISG - Brunnen]}$.

After that, we calculated the difference in distance per individual speaker (DD_speak) and lexical item. For each member of a dyad this is computed by subtracting d2 from d1 (DD_speak = d2 – d1). Negative values of DD_speak show evidence of convergence, positive values indicate divergence, values centred around zero signal maintenance.

To test (a), *i.e.*, whether one of the two dialects (ZHG) accommodates more and in which direction (convergence), we ran one Linear Mixed Effects Model with DD_speak as dependent variable and Dialect (GRG and ZHG) as fixed factor. The random part of the model comprised the intercept of Speakers and Lexical Items.

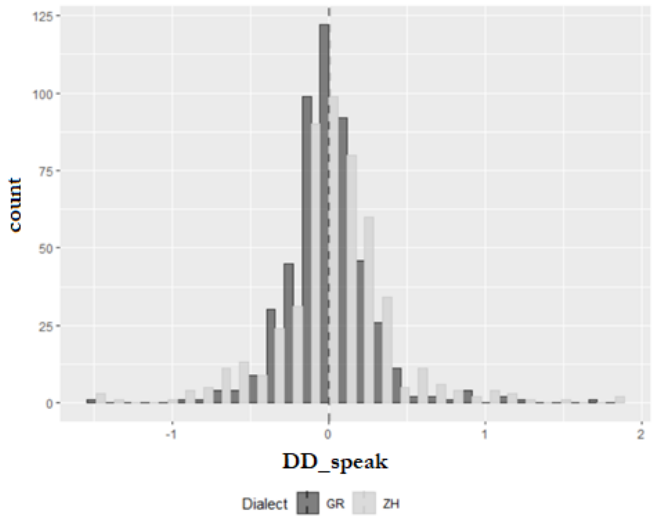
To test (b), *i.e.*, whether one of the two dialects accommodates more for the most acoustically distant ratio (RedVow), we tested the effect of the interaction between Dialect * Ratio Type (ISG, OSL, RedVow) on DD_speak, using a Linear Mixed Effects Model (Speakers and Lexical Items entered as random intercepts).

To test (c), *i.e.*, whether previous exposure to target lexical items during the diapix tasks evokes more convergence than novel words in one of the two dialects, we tested the effect of a three-way interaction between Dialect * Ratio Type * Word Type (novel vs previously heard items) on DD_speaker, with a Linear Mixed Effects Model (Speakers and Lexical items entered as random intercepts). Statistical analyses were performed with *R Studio* (R Core Team 2020).

2.3 Results and Discussion

Regarding (a), *i.e.,* whether ZHG speakers converged more towards GRG, the effect of Dialect on DD_speak was not significant [$x^2(1) = .0875$, $p = .767$)]. As shown in Fig. 2, the values of DD_speak largely overlap between the two dialects and the peaks of both groups are centred around zero. Against the predictions, ZHG speakers did not modify the production of the segmental timing properties towards their interlocutors' variety after dialogical interactions insomuch as they did for vowel qualities.

Figure 2 - *Histogram showing the distribution of DD_speak values within speakers per dialect*

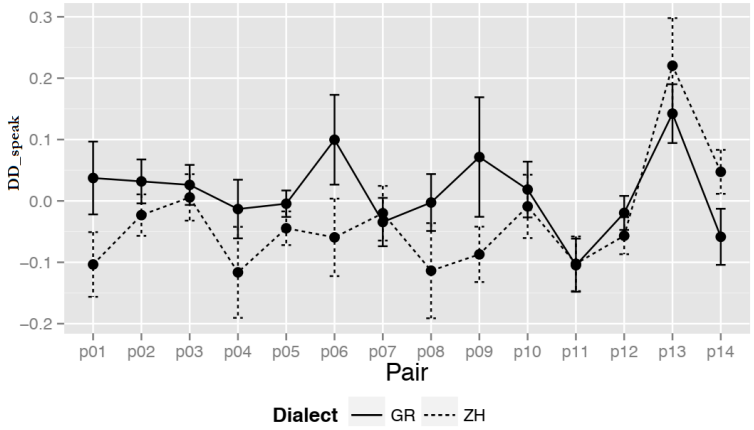


This is even more evident if we observe the DD_speak scores of ZHG and GRG speakers across pairs per vowel formants (data from Ruch, 2015) and the segmental durational properties examined in this study (Fig. 3a-b). The scores for vowel formants were mostly negative for ZHG (Fig. 3a – top panel), meaning that they were more inclined to converge, despite a remarkable variability between speakers

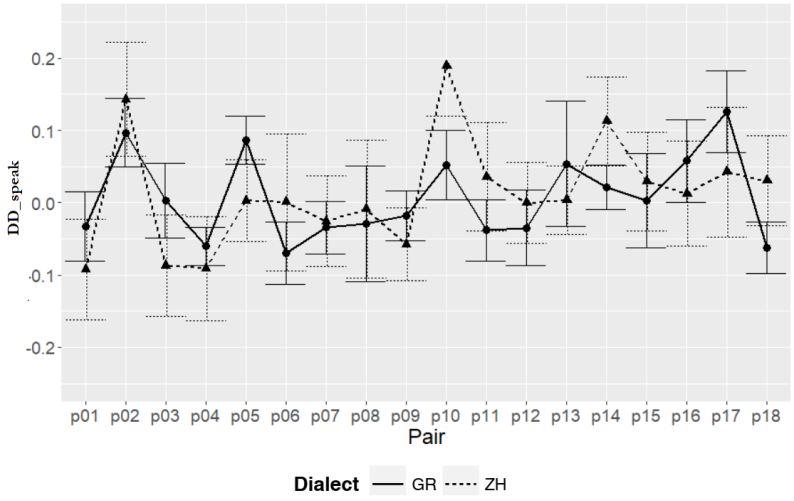
and pair (Ruch, 2015). For the segmental durational properties (Fig. 3b – bottom panel), instead, no clear pattern of convergence or divergence can be observed.

Figure 3a-b - *Line chart showing the difference in distance within speaker (DD_speak) per dialect and pair for vowel formants (top panel, plot in Ruch, 2015), and segmental durational properties (bottom panel). Dot represents average values, whisker error bar over all target words per speaker*

a) Vowel formants



b) Segmental durational properties

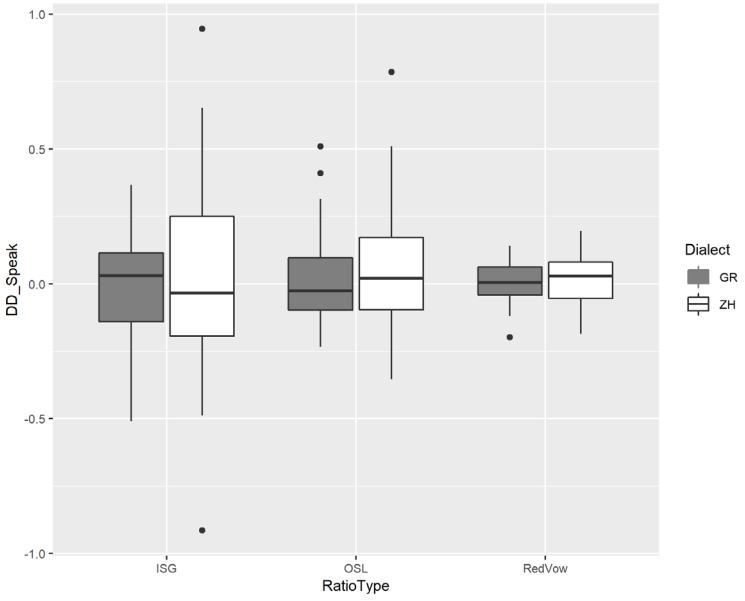


With respect to (b), *i.e.,* the hypothesis that ZHG and GRG are more prone to converge in RedVow as compared to OSL and ISG, the results of the statistical analysis have revealed that neither the interaction between Ratio Type and Dialect

[$x^2(2)= .422$, $p = .971$], nor the simple effects of Ratio Type and Dialect are significant [Ratio Type: $x^2(2)= 2.271$, $p = .678$; Dialect: $x^2(1)= .087$, $p = .767$].

Unlike other findings on cross-dialectal vowel accommodation showing more convergence for phonetically more distant vowels (e.g., Ruch, 2015; MacLeod, 2012), and more divergence for acoustic attributes perceived as strong dialect markers (e.g., Babel, 2010; Walker, Campbell-Kibler, 2015; Clopper, Dossey, 2020), none of the measures evoked a different degree of convergence in either dialect (Fig. 4). This was unexpected since the reduced/unreduced pronunciation of word final vowel (ɐ vs ə) is one of the features that best distinguishes the two dialects but is also a strong dialect marker. The same applies for OSL in comparison to ISG.

Figure 4 - *Difference in distance within speaker (DD_speak) per ratio types and dialects*

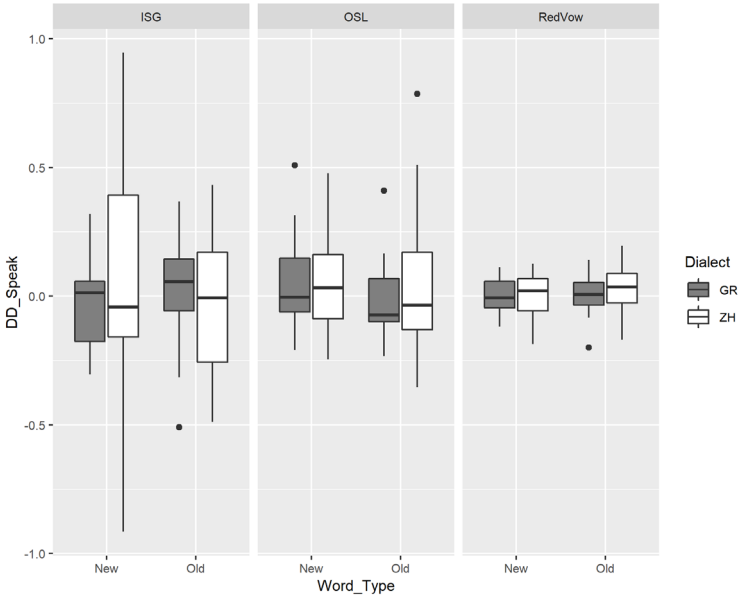


With respect to (c), *i.e.*, more accommodation is evoked in items which were previously used in the diapix tasks, ZHG and GRG speakers contradict the predictions. In the study on vowel convergence (Ruch, 2015), GRG diverged and ZHG converged more for low vowels in items which served as the basis for the diapix tasks. A similar effect of new *vs* old words on convergence was found also in imitation studies by Goldinger (1998) and Nielsen (2011). In the case of rhythmic accommodation, instead, ZHG and GRG speakers' inclination to rhythmically accommodate towards or away from their interlocutors' variety was not promoted nor inhibited by previous exposure to the items utilized in the diapix tasks (Fig. 5).

The three-way interaction between Dialect * Ratio Type * Word Type was not significant [$x^2(7)= 2.263$, $p = .943$)], nor were the two-way interactions between Word Type * Dialect [$x^2(1)= 3e\text{-}04$, $p = .986$], Word Type * Ratio Type [$x^2(2) = 1.352$,

p = .508] and Ratio Type * Dialect [$x^2(2) = .0424$, p = .979]. Insignificant were also the main effects of Dialect [$x^2(1) = .088$, p = .766], Ratio Type [$x^2(2) = 2.536$, p = .281] and Word Type [$x^2(1) = .701$, p = .402].

Figure 5 - *Difference in distance within speakers per ratio type, dialects and word type*



The findings of this study coupled with those in Pellegrino et al. (accepted) point to rhythmic maintenance. Both when analysed at the level of the pair and at the level of individual speakers, unexpectedly, the rhythmic distance between and within members of a dyad did not significantly vary from pre- to post-dialogue productions. Accommodation was not influenced by socio-linguistic factors (e.g., acoustic distance, dialect markedness) or by previous exposure to the target lexical items (previously heard vs new items).

One possible explanation for the difference in accommodation between vowel formants (Ruch, 2015) and the rhythmic features examined here can be related to the perceptual salience of prosodic and spectral cues for dialect identification. This line of research suggests that listeners can make use of a variety of cues – both temporal and spectral – to identify a dialect. Although, the prominence of certain cues over others varies noticeably from study to study (see a.o. Boula de Mareüil, Vieru-Dimulescu, 2006; Vicenik, 2011; Fuchs, 2015; Ruch, 2018), when it comes to Swiss German dialects, however, the pronunciation of vowels has been shown to play a major role as compared to prosodic cues (Leemann, Siebenhaar, 2008; Leemann, Kolly & Nolan, 2016; Leemann, Kolly, Nolan & Li, 2018; for varieties of English, cf. Fuchs, 2015). Although some of the features under examination (RedVow and OSL) were deemed to be salient features of ZHG and GRG (Ruch, 2018), vowel

pronunciation may have been more perceptually salient to interlocutors, and thus more susceptible to mutual adaptations between dialogue partners as the data in Ruch (2015) seem to suggest.

The results about rhythmic maintenance can also be explained along the same line of reasoning provided for the new data on cross-dialectal vowel accommodation (Ruch, 2021). The author imputed the conservative vowel production of ZHG and GRG speakers to the sociolinguistic situation of German Speaking Switzerland. As opposed to other countries, where dialects may be stigmatized in favour of standard variants, Swiss German dialects, instead, are employed in public and private domains, co-exist in everyday communicative situations due to their high mutual intelligibility, and share equal prestige (Scherrer et al. 2019). Although the two dialects vary in perceived likeability (Leemann et al. 2015; Ruch, 2018), in the absence of specific communication demands, ZHG and GRG speakers probably did not need to shift their pronunciation after being exposed to each other's dialect. Given that the speech produced during the diapix tasks was not object of the present investigation, we cannot, however, exclude that rhythmic adjustments between dialogue partners happened in the course of interactions but were not retained in post-dialogue productions.

There might be, however, explanations of different nature for the lack of a clear effect of speakers' dialect and word type on rhythmic accommodation. One for example may have to do with the small number of speakers (18 per dialect) and items (5 for ISG, 8 for OSL and 15 for RedVow). It is also plausible that by chance the pairing of GRG and ZHG speakers (at the corpus collection stage) resulted in 18 pairs with varying degree of baseline distance between the dialogue partners. This has been shown to be problematic for measurements of convergence based on the difference in distance approach (DD speak) used in research on phonetic convergence (Cohen Priva, Sanker, 2019). If the baseline distance is small, there is not much room for convergence, while the possibility of obtaining divergence is high. On the contrary, with large baseline distance, convergence can be overestimated. With a larger dataset, including data from the other tasks of the corpus, it would be interesting to see whether the same patterns of accommodation would replicate using the statistical model proposed by Cohen Priva, Sanker (2019). This overcomes the 'baseline bias' of the difference in distance approach, by taking the subject's and interlocutor's baselines as predictors of each subject's performance in mixed-effects linear regression.

Another factor which would deserve further investigation is the degree of within speakers' variability in the production of the examined rhythmic patterns prior to dialogical interactions. The results of a recent investigation examining computer simulations and experimental data from real interactions have, indeed, shown that the intrinsically more variable speaker of a dyad is the one who converges to the partner when accommodation takes place (Lee, Goldstein, Parrell & Byrd, 2021). Similarly, the results of the agent-based modeling of dialect contact reported in Ruch (2021) revealed that the dialect with the more dispersed distribution

converged towards the dialect with the more compact distribution. It would be thus interesting to test whether the speakers who maintain their rhythmic properties in post-dialogue productions are those with reduced acoustic variability in pre-dialogue recordings.

3. *Conclusions*

The results on rhythmic accommodation, together with findings on vowel convergence, based on the same corpus of lexical items in pre- and post-dialogue picture naming tasks, confirm the great complexity of the phenomenon of vocal accommodation. In line with findings showing that patterns of acoustic convergence are extremely variable within pairs or within speakers (Sanker, 2015; Cohen-Priva, Sanker, 2018), in this study the results on vowel convergence (Ruch, 2015) did not replicate for rhythmic accommodation. If ZHG speakers tended to converge more towards their dialect partner in terms of vowel formants, GRG and ZHG speakers tended to maintain their original rhythmic realization after being exposed to one another's dialect. Interpretations of accommodation based on phonetic distance and word type which were shown to apply for vowel formants do not, instead, hold for rhythmic patterns. Framing the results in the light of hypothesized effect of short-term accommodation on language variation and change (Trudgill, 1986), we can tentatively assume that vowel quality characteristics may play a major role in the diffusion of linguistic innovations and dialectal levelling in German speaking Switzerland. Under the given experimental conditions, segmental timing features, indeed, were less prone to accommodation as opposed to vowel formants. In view of the more recent data on vowel maintenance (Ruch, 2021), these assumptions will warrant further research based on the more spontaneous tasks of the corpus, or using data collected in more ecologically valid communicative situations. Our data are based on isolated lexical items from a picture naming task without any pragmatic context, and this is certainly a type of speech that typically does not play a large role in everyday communication.

Acknowledgments

This research was supported by the University Research Priority Program (URPP) "Language and Space" of the University of Zurich (UZH). The author would like to thank Hanna Ruch for the corpus, Sandra Schwab for the feedback on the statistical procedure, and the reviewers for their helpful comments on an earlier version of this manuscript.

Bibliography

Babel, M. (2010). Dialect divergence and convergence in New Zealand English. In *Language in Society*, 39(4), 437-456. https://doi:10.1017/S0047404510000400.

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. In *Journal of Phonetics*, 40, 117-189. https://doi.org/10.1016/j.wocn.2011.09.001.

Boula de Mareüil, P., Vieru-Dimulescu, B. (2006). The contribution of prosody to the perception of foreign accent. *Phonetica,* 63(4), 247-267. 10.1159/000097308.

Brennan, S.E. (1996). Lexical entrainment in spontaneous dialog. *Proceedings of the International Symposium on Spoken Dialogue*, Philadelphia, PA, 41-44.

Bullock, B., Gerfen, C. (2004). Phonological convergence in a contracting language variety. In *Bilingualism: Language and Cognition*, 7(2), 95-104. 10.1017/S1366728904001452.

Carli, L.L., LaFleur, S.J. & Loeber, C.C. (1995). Nonverbal behavior, gender, and influence. In *Journal of Personality and Social Psychology*, 68(6), 1030-1041. https://doi.org/10.1037/0022-3514.68.6.1030.

Cerda-Oñate, K., Toledo Vega, G. & Ordin, M. (2021). Speech rhythm convergence in a dyadic reading task. In *Speech Communication*, 131, 1-12. https://doi.org/10.1016/j.specom.2021.04.003.

Chartrand, T.L., Bargh, J.A. (1999). The chameleon effect: the perception-behavior link and social interaction. In *Journal of Personality and Social Psychology,* 76 (6), 893-910. https://doi.org/10.1037/0022-3514.76.6.893.

Clopper, C.G., Dossey, E. (2019). Phonetic convergence and divergence in the American Midwest. In *The Journal of the Acoustical Society of America*, 145(3), 1930-1931.

Clopper, C.G., Dossey, E. (2020). Phonetic convergence to Southern American English: Acoustics and Perception. In *The Journal of the Acoustical Society of America*, 147(1), 671-683. https://doi.org/10.1121/10.0000555.

Cohen Priva, U., Sanker, C. (2018). Distinct behaviors in convergence across measures. *Annual Conference of the Cognitive Science Society*, Madison, WI, 1518-1523.

Cohen Priva, U., Sanker, C., (2019) Limitations of difference-in-difference for measuring convergence. *Laboratory Phonology* 10(1), 15, 1-29. https://doi.org/10.5334/labphon.200.

D'Imperio, M., Cavone, R. & Petrone, C. (2014). Phonetic and phonological imitation of intonation in two varieties of Italian. In *Frontiers in Psychology*, 5:1226. https://doi.org/10.3389/fpsyg.2014.01226.

Eckhardt, O. (1991). Die Mundart der Stadt Chur. Zürich: Phonogrammarchiv der Universität Zürich.

Edlund, J., Heldner, M. & Hirschberg, J. (2009). Pause and gap length in face-to-face interaction. *10th Annual Conference of the International Speech Communication Association*, 2779-2782.

Evans, B.G., Iverson, P. (2007). Plasticity in vowel perception and production: a study of accent change in young adults. In *Journal of the Acoustical Society of America*, 121, 3814-3826. 10.1121/1.2722209.

Ferguson, C.A. (1975). Toward a characterization of English foreigner talk. In *Anthropological Linguistics*, 17(1), 1-14. https://doi.org/10.2307/30027270.

Fleischer, J., Schmid, S. (2006). Zurich German. In *Journal of the International Phonetic Association*, 36, 243–253. 10.1017/S0025100306002441.

Fuchs, R. (2015). You're not from around here, are you? In Delais-Roussarie, E., Avanzi, M. & Herment, S. (Eds.) *Prosody and Language in Contact*. Berlin: Springer, 123-148.

Giles, H. (2008). Communication accommodation theory. In Baxter, L.A., Braithwaite, D.O. (Eds.). *Engaging theories in interpersonal communication: Multiple perspectives.* Thousand Oaks: Sage Publications, 161-173.

Giles, H., Mulac, A., Bradac, J.J. & Johnson, P. (1987). Speech Accommodation Theory: The first decade and beyond. In *Annals of the International Communication Association*, 10(1), 13-48. https://doi.org/10.1080/23808985.1987.11678638.

Goldinger, S.D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105(2), 251-279. 10.1037/0033-295x.105.2.251.

Hughes, S.M., Mogilski, J.K. & Harrison, M.A. (2014). The perception and parameters of intentional voice manipulation. In *Journal of Nonverbal Behavior*, 38(1), 107-127. https://doi.org/10.1007/s10919-013-0163-z.

Kemper, S., Finter-Urczyk, A., Ferrell, P., Harden, T. & Billington, C. (1998). Using elderspeak with older adults. In *Discourse Processes*, 25(1), 55-73. https://doi.org/10.1080/01638539809545020.

Kuhl, P.K., Andruski, J.E., Chistovich, I.A., Chistovich, L.A., Kozhevnikova, E.V., Ryskina, V.L., Stolyarova, E.I., Sundberg, U. & Lacerda, F. (1997). Cross-Language analysis of phonetic units in language addressed to infants. In *Science*, 684-686. 10.1126/science.277.5326.684.

Lee, Y., Goldstein, L., Parrell, B. & Byrd, D. (2021). Who converges? Variation reveals individual speaker adaptability. In *Speech Communication*, 131, 23-34. https://doi.org/10.1016/j.specom.2021.05.001.

Leemann, A. (2012). *Swiss German Intonation Patterns*. Amsterdam: John Benjamins Publishing Company.

Leemann, A., Dellwo, V., Kolly, M.-J. & Schmid, S. (2012). Rhythmic variability in Swiss German dialects. *Proceedings of the 6th International Conference on Speech Prosody*, Shanghai, China, 22-25 May 2012, 607-610.

Leemann, A., Kolly, M.-J., & Nolan, F. (2015). It's not phonetic aesthetics that drives dialect preference: The case of Swiss German. *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, UK, 10-14 August 2015.

Leemann, A., Kolly, M.-J. & Nolan, F. (2016). Identifying a speaker's regional origin: The role of temporal information. *Proceedings of Speech Prosody*, Boston, MA, 31 May-3 June 2016.

Leemann, A., Kolly, M.-J., Nolan, F. & Li, Y. (2018). The role of segments and prosody in the identification of a speaker's dialect. In *Journal of Phonetics*, 68, 69-84. https://doi.org/10.1016/j.wocn.2018.02.001.

Leemann, A., Siebenhaar, B. (2008). Perception of dialectal prosody. *Proceedings of the 9th Annual Conference of the International Speech Communication Association*, Brisbane, Australia, 22-26 September 2008, 524-527.

Lelong, A., Bailly, G. (2010). Study of the phenomenon of phonetic convergence thanks to speech dominoes. *Proceedings of Analysis of Verbal and Nonverbal Communication and Enactment. The Processing Issues – COST 2102 International Conference*, Budapest, Hungary, September 7-10, 273-286.

Leong V., Kalashnikova, M., Burnham, D. & Goswami, U. (2017). The temporal modulation structure of infant-directed speech. In *Open Mind: Discoveries in Cognitive Science,* 1, 78-90. https://doi.org/10.1162/OPMI_a_00008.

MacLeod, B. (2012). The effect of perceptual salience on phonetic accommodation in cross-dialectal conversation in Spanish. PhD Thesis, University of Toronto.

Mitterer, H., Müsseler, J. (2013). Regional accent variation in the shadowing task: Evidence for a loose perception-action coupling in speech. In *Attention, Perception & Psychophysics*, 75, 557-575. https://doi.org/10.3758/s13414-012-0407-8.

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. In *Journal of Phonetics*, 39(2), 132-142. https://doi.org/10.1016/j.wocn.2010.12.007.

Nguyen, N., Delvaux, V. (2015). Role of imitation in the emergence of phonological systems. In *Journal of Phonetics*, 53, 46-54. https://doi.org/10.1016/j.wocn.2015.08.004.

Pardo, J. (2012). Reflections on phonetic convergence: Speech perception does not mirror speech production. In *Language and Linguistics Compass*, 753-767. https://doi.org/10.1002/lnc3.367.

Pardo, J.S., Urmanche, A., Wilman, S., & Wiener, J. (2017). Phonetic convergence across multiple measures and model talkers. In *Attention, Perception, & Psychophysics*, 79(2), 637-659. https://doi.org/10.3758/s13414-016-1226-0.

Payne, E., Post, B., Astruc, L., Prieto, P. & Vanrell, M. d. M. (2009). Rhythmic modification in child directed speech. In *Working Papers in Linguistics, Philology & Phonetics*, 12, 123-144. http://www.ling-phil.ox.ac.uk/files/uploads/OWP2009.pdf.

Pellegrino, E., Schwab S. & Dellwo, V. (accepted). Do speakers converge rhythmically? A study on Grison and Zurich German, *Loquens.*

Pentland, A. (2008). *Honest Signal: How they shape our world*. Cambridge: MIT Press.

Pickering, M.J., Garrod, S. (2004). Toward a mechanistic psychology of dialogue. In *Behavioral & Brain Sciences* 27, 169-190. 10.1017/S0140525X04000056.

Pickering, M.J., Garrod, S. (2006). Alignment as the basis for successful communication. In *Research on Language and Computation* 4 (2–3), 203-228. 10.1007/s11168-006-9004-0.

R Core Team. (2020). A language and environment for statistical computing. [Computer Program] Version 4.0.1, retrieved 6 June 2020 from https://www.r-project.org/.

Romera, M., Elordieta, G. (2013). Prosodic accommodation in language contact: Spanish intonation in Majorca. In *International Journal of the Sociology of Language*, vol. 2013, no. 221, 127-151. https://doi.org/10.1515/ijsl-2013-0026.

Ross, J., Lilley, K., Clopper C., Pardo J. & Levi, S. (2021). Effects of dialect-specific features and familiarity on cross-dialect phonetic convergence. In *Journal of Phonetics*, 86, 101041, 1-23. https://doi.org/10.1016/j.wocn.2021.101041.

Ruch, H. (2015). Vowel convergence and divergence between two Swiss German dialects. *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, UK, 10-14 August 2015.

Ruch, H. (2018). The role of acoustic distance and sociolinguistic knowledge in dialect identification. In *Frontiers in Psychology*, 9:818. https://doi.org/10.3389/fpsyg.2018.00818.

Ruch, H. (2021). Dialect contact in real interaction and in an agent-based model. In *Speech Communication*, 134, 55-70. https://doi.org/10.1016/j.specom.2021.09.003.

Ruch, H., Zürcher, Y. & Burkart, J.M. (2017). The function and mechanism of vocal accommodation in humans and other primates. In *Biol Rev Camb Philos Soc*. 2018 May, 93(2), 996-1013. 10.1111/brv.12382.

Sanker, C. (2015). Comparison of phonetic convergence in multiple measures. In *Cornell Working Papers in Phonetics and Phonology 2015,* 60-75.

Scherrer, Y., Samardžić, T. & Glaser, E. (2019). Digitising Swiss German: How to process and study a polycentric spoken language. In *Language Resources and Evaluation*, 53, 735-769. https://doi.org/10.1007/s10579-019-09457-5.

Summers, W.V., Pisoni, D.B., Bernacki, R.H., Pedlow, R.I. & Stokes, M.A. (1988). Effects of noise on speech production: Acoustic and perceptual analyses. In *The Journal of the Acoustical Society of America*, 84(3), 917-928. 10.1121/1.396660.

Traunmüller, H., Eriksson, A. (2000). Acoustic effects of variation in vocal effort by men, women, and children. In *The Journal of the Acoustical Society of America*, 107(6), 3438-3451. 10.1121/1.429414.

Troncoso-Ruiz, A., Elordieta, G. (2017). Prosodic accommodation and salience: The nuclear contours of Andalusian Spanish speakers in Asturias. In *Loquens*, 4(2), e043. https://doi.org/10.3989/loquens.2017.043.

Trudgill, P. (1986). *Dialects in Contact*. Oxford: Blackwell Publishing.

Van Engen, K.J., Baese-Berk, M., Baker, R.E., Choi, A., Kim, M. & Bradlow, A.R. (2010). The Wildcat Corpus of native-and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. In *Language and Speech*, 53(4), 510-540. 10.1177/0023830910372495.

Vicenik, C.J. (2011). The role of intonation in language discrimination by infants and adults. PhD Thesis, University of California, Los Angeles.

Walker, A., Campbell-Kibler, K. (2015). Repeat what after whom? Exploring variable selectivity in a cross-dialectal shadowing task. In *Frontiers in Psychology*, 6, 546. https://doi.org/10.3389/fpsyg.2015.00546.

Werlen, I., Schlegel, D. (2006). *Zwischen 'Grüessech' und 'Tagwoll'. Das Sprachverhalten und die Lebenssituation der Oberwalliser und Oberwalliserinnen in Bern*. Institut für Sprachwissenschaft, Universität Bern, Bern.