Syllable rate and speech rhythm in multiethnolectal Zurich German: A comparison of speaking styles

Morand, Marie-Anne ; Bruno, Melissa ; Schwab, Sandra ; Schmid, Stephan

DOI: https://doi.org/10.21437/SpeechProsody.2022-69

Speech Prosody 2022
23-26 May 2022, Lisbon, Portugal

Syllable rate and speech rhythm in multiethnolectal Zurich German: A comparison of speaking styles

Marie-Anne Morand[1], *Melissa Bruno*[1], *Sandra Schwab*[2], *Stephan Schmid*[1]

[1]University of Zurich, Switzerland
[2]University of Bern, Switzerland
{marie-anne.morand, melissa.bruno, stephan.schmid}@uzh.ch,
sandra.schwab@unibe.ch

Abstract

Multiethnolectal ways of speaking have been emerging for 30 years in culturally and linguistically diverse neighborhoods of European cities, including Zurich (Switzerland). Among the prosodic features of Germanic multiethnolects, a so-called 'staccato' rhythm has been mentioned in several studies. For instance, a comparison between two groups of adolescents (12 speakers each) showed that speakers of multiethnolectal Zurich German displayed slower syllable rates and less vowel duration variability than speakers of a rather traditional dialect.

This study compares syllable rate and speech rhythm metrics (*nPVI-V, nPVI-C*) in spontaneous and read speech of 48 Zurich German adolescents. In a regression analysis, rhythmic measures were compared with the perception of how multiethnolectal the speakers sounded (*rating score*). The results showed that syllable rate and *nPVI-V* were related to rating score independently of speaking style (read, spontaneous speech): Speakers who were perceived as more multiethnolectal had a slower syllable rate and less vowel duration variability. Such findings were not observed for *nPVI-C.*

These results suggest that syllable rate and speech rhythm (at least, vowel duration variability) are stable phonetic features of multiethnolectal Zurich German, since the relationship between these features and the perception of multiethnolectal speech was observed in both read and spontaneous speech.

Index Terms: speech rhythm, multiethnolects, speaking style, Zurich German, sociophonetics

1. Introduction

In many linguistically and culturally diverse neighborhoods of European cities, so-called 'multiethnolectal' [1] ways of speaking have been emerging at least since the 1990s. Outside Europe, ethnolects have been studied, for instance, in Toronto [2]. Along with grammatical and lexical innovations, such ways of speaking also differ from more traditional vernaculars of the respective cities with respect to various segmental and suprasegmental features. Speech rhythm is one of these phonetic features. For several Germanic languages, it has been noticed that multiethnolectal varieties display, for example, less vowel duration variability which leads to the impression of a so-called 'staccato' rhythm (e.g., Danish [3], English [4], or Swedish [5]).

As is well known, there are different phonological and phonetic approaches to the study of speech rhythm. Pike [6] and Abercrombie [7] established the original 'isochrony hypothesis', according to which, most of the world's languages can be classified either as 'stress-timed' or as 'syllable-timed'; later, a 'mora-timed' rhythm type was postulated as well (e.g., [8]). However, research in experimental phonetics has found no evidence in support of the classical isochrony hypothesis (see [9] and [10] for an overview). Consequently, new solutions to the problem of linguistic rhythm were proposed. From a phonological perspective, different languages have been categorized as either 'syllable-based' or 'word-based' [11], [12], which represent two ends of a continuum [9], [13]. Word languages are claimed to have rather complex syllable structures and a durational reduction of non-accented syllables, whereas syllable languages typically have simple syllable structures and no accent-dependent reduction [11, p. 11], amongst other features.

At the turn of the millennium, new 'rhythm metrics' were proposed to find acoustic correlates of the phonological 'rhythm classes' (see [13] for an overview). These rhythm metrics were based on the durations of so-called vocalic and consonantal intervals, more precisely on their standard deviations (as suggested by Ramus [14]) or their variation coefficients (as proposed by Dellwo [15]). Additionally, Grabe and Low [16] suggested the raw and normalized Pairwise Variability Indices (*rPVI*, *nPVI*), calculated as the mean of the duration differences of all subsequent vocalic and consonantal intervals.

The original goal of such rhythm metrics was to compare different languages (assuming that the temporal structure of vocalic and consonantal intervals would reflect phonological properties of these languages [17], [18]). However, some scholars found differences in speech rhythm, even between two varieties of the same language (e.g., between European and Brazilian Portuguese [19]) or between L1 and L2 speakers [20]. Rhythm metrics were even applied to capture individual differences in the temporal organization of speech [21].

It must be considered, however, that severe criticism against the usefulness of rhythm metrics has been raised as well (e.g., [22]). Many rhythm studies have analyzed read speech, but even with carefully designed sentences the phonotactic makeup of the read words can have a notable impact on rhythm metrics. Leemann and colleagues [30] compared rhythm metrics obtained for eight Swiss German dialects in three different sentences to test the effect of the sentences used. They found a significant effect of sentence for most investigated rhythm metrics as well as a significant interaction for sentence and dialect. Thus, their results demonstrate that sentence selection is crucial in speech rhythm research.

Fewer studies investigated spontaneous speech, probably because it can be full of hesitations, repairs, and longer pauses. Therefore, one must be careful when interpreting results based on rhythm metrics and always consider the possible factors that may determine the structure of the data (§3).

The present study investigates temporal characteristics (rate and rhythm) in one single dialect, Zurich German, analyzing

10.21437/SpeechProsody.2022-69

read and spontaneous speech of 48 Zurich German adolescents (see §4.1). It is hypothesized that temporal characteristics act as a sociophonetic marker [23] across speaking styles and that rhythmic differences are mainly due to typical duration patterns of stressed and unstressed vowels.

In a previous study on read speech, we found slower syllable rates in speakers of rather multiethnolectal Zurich German (MEZ) compared to speakers of rather traditional Zurich German [24]. Based on a rating experiment, adolescents could be placed on a perceptual continuum between multiethnolectal and traditional Zurich German [25]. It now remains to be seen whether syllable rate also proves to be linked to multiethnolectality in spontaneous speech and thus could be classified as a general prosodic feature of MEZ.

Our previous study found less durational variability of vocalic intervals in speakers who were rated as speaking more multiethnolectal Zurich German. The current study will therefore investigate the extent to which this also applies to spontaneous speech (or maybe even more so; see below).

In the following, we will first describe rhythmic features of traditional Zurich and Swiss German before reviewing results on multiethnolectal varieties of other Germanic varieties. Then, the analysis of syllable rate and speech rhythm in spontaneous speech of adolescents from Zurich will be presented and compared to the previous results of read speech from the same adolescents. The comparison will be based on the *nPVI* for vocalic and consonantal intervals in spontaneous speech as well as on syllable rate measured in syllables per second.

2. Rhythm in Zurich German

2.1. Traditional Zurich German

Zurich German is a High Alemannic dialect spoken in and around the city of Zurich (Switzerland). In the light of Auer's [11] phonological typology, Nübling and Schrambke [26] argue that Alemannic dialects exhibit some characteristics of syllable languages. This view is partly shared by Siebenhaar [27] who nevertheless also points to phonotactic properties of Swiss German dialects that are typical of word languages. Indeed, acoustic investigations applying rhythm metrics [28], [29], [30], [31] do place Zurich German rather among the languages that have traditionally been considered as stress-timed.

2.2. Multiethnolectal Zurich German

Whereas traditional dialects and standard varieties of many Germanic languages are usually regarded as stress-timed, it has been noticed that multiethnolectal speaking styles show temporal features of rather syllable-timed languages [3], [4], [5]. These features include less durational variability of vocalic intervals measured as lower values for the *nPVI-V*.

For MEZ, a preliminary study [32] found a slower articulation rate (segments per second) for a multiethnolectal speaker (13.25 seg/sec) compared to another young speaker who spoke a rather traditional dialect (14.22 seg/sec). This characteristic of MEZ is taken up by a comedian to caricature the multiethnolectal way of speaking in a very exaggerated way (11.24 seg/sec [32]). Subsequently, our first quantitative analysis of read speech from 48 adolescents recorded in Zurich [24] revealed significant negative correlations between the perceptual ratings assigned to multiethnolectal speakers and their syllable rate as well as their vowel duration variability. Thus, speakers who were perceived as speaking rather MEZ showed slower syllable rates and less vowel duration variability.

There is reason to believe that multiethnolectal features are more pronounced in spontaneous speech than in read speech, as the communicative situation may have an impact on the use of certain features. Wiese and Pohle [33], for example, found a certain syntactical feature of multiethnolectal German significantly more often in informal peer-group situations than in formal contexts. However, a preliminary analysis [34] of a subset of 24 out of the 48 speakers involved in the current contribution showed similar trends both in read and spontaneous speech.

3. Rhythm in read and spontaneous speech

The current study investigates syllable rate and speech rhythm of adolescents in two different speaking styles (i.e., read and spontaneous speech). As a matter of fact, results yielded by rhythmic metrics may be influenced not only by the phonotactic structures contained in the analyzed sentence material (e.g., [22], [30]), but also by speaking style.

For instance, Arvaniti [22] compared three speaking styles (spontaneous speech, story, and sentence reading) for six different languages with eight speakers each and found that traditional rhythm classifications were inconsistent across several rhythm metrics (*%V*, *ΔC*, *nPVI*, *rPVI*, *VarcoC*, and *VarcoV*), which showed much inter-speaker variation and proved very sensitive to speaking style and syllable complexity.

3.1. Research questions

Based on the literature review, we have formulated the following four research questions:

RQ1: To what extent are slower syllable rates related to MEZ in spontaneous speech?

RQ2: To what extent is less durational variability of consonantal intervals related to MEZ in spontaneous speech?

RQ3: To what extent is less durational variability of vocalic intervals related to MEZ in spontaneous speech?

RQ4: To what extent does speaking style influence syllable rate and speech rhythm in MEZ?

For Zurich German, we assume to find a relation between the degree of perceived multiethnolectality of the speakers on the one hand and their syllable rate (RQ1) and the vowel duration variability on the other hand (RQ3). However, we do not expect a relation between the degree of perceived multiethnolectality and the consonant duration variability (RQ2). There are no clear predictions regarding RQ4; we assume, however, that multiethnolectal features might be more pronounced in spontaneous speech than in read speech.

4. Data and methods

4.1. Speakers

Forty-eight adolescents were recorded at two schools in the city of Zurich (28 females; mean age = 14.27 years; *SD* = 0.74). Both schools are in a working-class neighborhood. Most adolescents (n = 39) spoke additional or other languages than Zurich German before they entered kindergarten. First languages include – but are not limited to – the following (in alphabetical order): Albanian (n = 8), Arabic (n = 3), French (n = 2), Portuguese (n = 3), Punjabi (n = 1), Somali (n = 1),

Spanish (n = 4), and Turkish (n = 2). In many cases (n = 31), both parents are citizens of a country other than Switzerland.

4.2. Material

To elicit spontaneous speech, adolescents played a game of 'spot the difference' in pairs using so-called *Diapix* [35], [36], which resulted in recordings of 10 to 20 minutes. Utterances of at least three syllables were selected, which yielded a total of 6068 utterances. The number of utterances per speaker varies between 20 and 108 (mean = 65). Utterance length varies between 3 and 29 syllables (mean = 7.2 syllables/utterance).

The data analyzed for read speech was recorded in a battery of carefully designed sentences, which contained 3 to 7 syllables. In total, speakers read 100 sentences presented individually on a computer screen in random order. In total, 2949 sentences were analyzed. Per speaker, 37-84 sentences were analyzed (mean = 61).

4.3. Rating scores

To obtain an 'emic' perspective [37] – i.e., as from inside the cultural system of MEZ – we conducted a 'rating experiment' with a different, but very similar group of adolescents [25]. Speech samples of the 48 speakers were rated by 40 pupils (25 females, mean age = 14.8 years, range = 14-16 years) from another Zurich school. The samples of 5 to 7 seconds were extracted from the middle of a picture description task. Almost half of the raters (n = 19) are bilinguals. The raters' (additional) first languages include – but are not limited to – the following (in descending order): Tamil (n = 3), Turkish (n = 3), English (n = 2), Kurdish (n = 2), Italian (n = 2), and Spanish (n = 2). Raters had to evaluate on a 7-point Likert scale how multiethnolectal the speakers sounded ('To what extent do you think the person speaks slang ("German of foreigners")?'; 1 = not at all, 7 = completely). Rating scores correspond to the mean values of all 40 raters.

4.4. Data analysis

To analyze syllable rate and speech rhythm in read and spontaneous speech, we transcribed all stereo recordings in *Praat* [38] using the spelling conventions established by Dieth [39]. All read and spontaneous utterances were first segmented automatically with *WebMAUS* [40] at the level of words, syllables, and phones; afterwards, segmentation was checked manually. Hesitations (e.g., *uh*, *uhm*) were excluded from the calculations.

Based on the syllable and phone segmentation, we measured the duration for vocalic (v) and consonantal (c) intervals to calculate *nPVI-V* and *nPVI-C* as well as speech rate, using the *Praat* plugin 'Duration Analyzer' [41].

To answer the research questions formulated in §3.1, we conducted different statistical analyses. First, we correlated rhythm metrics as well as syllable rate for read and spontaneous speech (§5.1). Afterwards, generalized estimation equations (GEE) were fitted to examine rating scores as a function of syllable rate and speaking style (§5.2) as well as a function of the variability of the duration of consonantal and vocalic intervals (§5.3, §5.4) and speaking style. GEE allows for the analysis of repeated measurements like those observed in our study: Speakers produce several utterances from which we extracted measurements. These data are thus speaker-dependent. Rating score was the dependent variable in all models. Independent variables of all our initial models included temporal metrics (syllable rate, *nPVI-C*, and *nPVI-V*), speaking style (read and spontaneous speech) as well as the interaction of the two factors. Since none of the interactions were significant, they are not included in the final models.

5. Results

5.1. Comparison of read and spontaneous speech

As can be observed in Table 1, there are significant positive correlations between read and spontaneous speech for all three measurements (syllable rate, *nPVI-C*, and *nPVI-V*). Speakers with a fast syllable rate and high durational variability in read speech tend to have a fast syllable rate and high durational variability also in spontaneous speech.

Table 1: *Pearson correlations between read and spontaneous speech.*

Measurement	Pearson *r*	*p*-value
Syllable rate	.704	< .001
nPVI-C	.421	.003
nPVI-V	.400	.005

5.2. Syllable rate

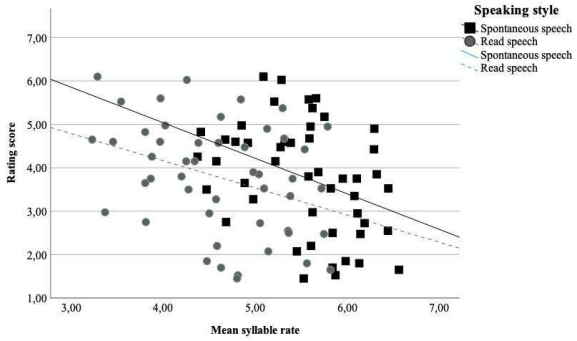

Figure 1: *Rating score as a function of mean syllable rate in read speech (grey points, dashed line) and spontaneous speech (black squares, solid line).*

As becomes evident from Figure 1, rating score and mean syllable rate are related in a similar way in read speech (R^2 = .125) and in spontaneous speech (R^2 = .143). In general, adolescents who were rated as speaking rather multiethnolectal Zurich German read and spoke slower.

The statistical analysis showed a significant effect of syllable rate (Wald χ^2 (1) = 12.261, p < .001) as well as of speaking style (Wald χ^2 (1) = 10.443, p = .001) on rating scores. In general, adolescents who read and spoke slower displayed higher rating scores (i.e., were rated as speaking more multiethnolectal Zurich German). Thus, it is safe to assume that syllable rate may influence the degree of perceived multiethnolectality not only in spontaneous, but also in read speech.

5.3. Variability of the duration of consonantal intervals

Figure 2 displays rating score as a function of the mean *nPVI* for consonantal intervals (*nPVI-C*) per speaker in read (R^2 = .018) and spontaneous speech (R^2 = .001). Although there is a significant correlation between speaking styles (see Table 1), there seems to be no significant relation between the consonant duration variability and adolescents' rating scores. This observation is confirmed by the statistical analysis, as there was no significant effect of *nPVI-C* (Wald χ^2 (1) = .305,

p = .581) on the degree of perceived multiethnolectality (rating score).

On average, durational variability of consonantal intervals is slightly higher in spontaneous speech (55.11, *SD* = 4.21) than in read speech (51.28, *SD* = 3.06). However, there was no significant effect of speaking style (Wald χ^2 (1) = .303, p = .582).

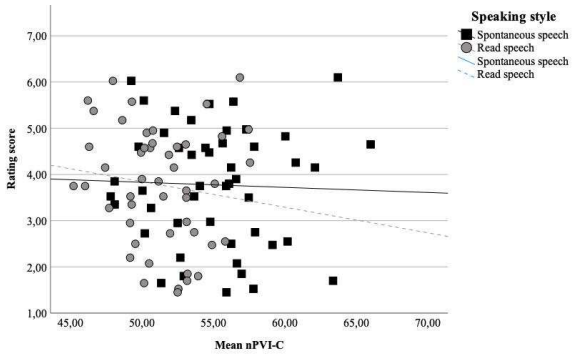

Figure 2: *Rating score as a function of mean* nPVI-C *in read speech (grey points, dashed line) and spontaneous speech (black squares, solid line).*

5.4. Variability of the duration of vocalic intervals

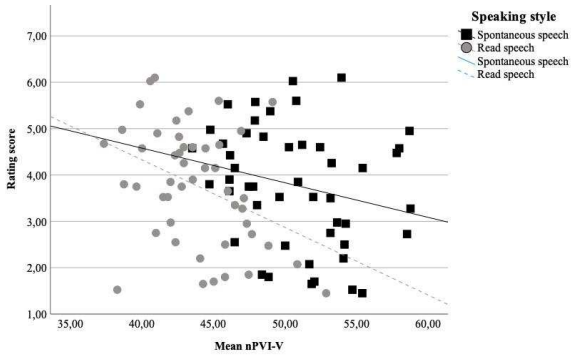

Figure 3: *Rating score as a function of mean* nPVI-V *in read speech (grey points, dashed line) and spontaneous speech (black squares, solid line).*

Lastly, Figure 3 shows how rating score is related to the *nPVI* for the duration of vocalic intervals (*nPVI-V*) both in read (R^2 = .145) and spontaneous speech (R^2 = .056). It appears that more durational variation is found in adolescents who were rated as speaking rather traditional Zurich German (i.e., low rating scores) than in adolescents who were rated as speaking rather multiethnolectal Zurich German. The statistical analysis confirmed this observation. There was a significant effect of *nPVI-V* (Wald χ^2 (1) = 9.324, p = .002) on the degree of perceived multiethnolectality (rating score).

Moreover, all adolescents showed more durational variability of vocalic intervals in spontaneous speech (50.78, *SD* = 4.07) than in read speech (43.80, *SD* = 3.33). This is also confirmed by the statistical analysis as we find a significant effect of speaking style (Wald χ^2 (1) = 10.320, p = .001).

6. Discussion and conclusion

We observe positive correlations of different temporal metrics between read and spontaneous speech. As already mentioned, studies on linguistic rhythm are usually based on read speech, where sentence material can be controlled. Therefore, it is worth noting that in our data the observed effects in read speech are highly correlated with the observed effects in spontaneous speech. What is more crucial, we did not find any significant interactions of temporal metrics and speaking style. Although syllable rate and variability of the duration of vocalic intervals were higher in spontaneous speech, they were systematically higher for all adolescents regardless of their rating scores.

As mentioned before (§2.2), it seemed that slower syllable rate could be a potential suprasegmental feature of MEZ. A speaker of traditional Zurich German showed a faster articulation rate in a preliminary study [32] than a speaker of MEZ. The more robust data of the present study now show a more detailed picture. It is true that adolescents who have been rated as speaking very traditional Zurich German generally tend to have a faster syllable rate, whereas multiethnolectal speakers are more evenly distributed in that they may have both slow and fast syllable rates.

Regarding *nPVI-C,* Leemann et al. [29, p. 609] found high variation between eight Swiss German dialects; in addition, Zurich German demonstrated the biggest range for this rhythm metric (range ca. 30-90, approximated from their figure). In our data, we also find some variation for *nPVI-C* (range for read speech 45.22-57.56; range for spontaneous speech 47.82-65.98). However, the degree of perceived multiethnolectality does not seem to be related to this rhythmic metric.

Instead, a clear finding of the current study indicates that adolescents who were perceived as speaking more MEZ displayed less durational variability of vocalic intervals both in read and spontaneous speech. This appears to be indeed a constant multiethnolectal feature across speaking styles in Zurich German as we did not find a significant interaction of speaking style and *nPVI-V* in our model. Arguably, it might be the case that multiethnolectal phonetic features might be less controllable than syntactic structures (see [33]) and therefore less prone to vary with respect to communicative context.

In sum, the results of the present study demonstrate that low vowel duration variability is a stable phonetic feature of MEZ, as the relation between this metric and the perception of multiethnolectal speech was observed in read as well as in spontaneous speech. With respect to syllable rate, the picture appears to be less clear. Rather than a slower syllable rate being a stable phonetic feature of MEZ, faster syllable rates seem to be characteristic of traditional Zurich German.

Regarding the more general issue of speech rhythm and multiethnolects, our study provides further evidence for the role of vowel duration variability as a sociophonetic marker in Germanic languages (§1). This might be the case not only of multiethnolectal Swiss German, but also multiethnolects spoken in Germany as well.

7. Acknowledgements

This research was supported by the Swiss National Science foundation (Grant No. 165798). We would like to thank all the speakers, raters, and their teachers for the participation in this study.

8. References

[1] M. Clyne, “Lingua Franca and ethnolects in Europe and beyond,” *Sociolinguistica*, vol. 14, pp. 83–89, 2000.

[2] M. F. Hofmann and J. A. Walker, “Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English,” *Language Variation and Change*, vol. 22, pp. 37–67, 2010.

[3] G. F. Hansen and N. Pharao, “Prosody in the Copenhagen Multiethnolect,” in P. Quist and B. A. Svendsen (eds.), *Multilingual Urban Scandinavia: New Linguistic Practices*. Bristol: Multilingual Matters, pp. 79–95, 2010.

[4] E. Torgersen and A. Szakay, “A study of rhythm in London: Is syllable-timing a feature of Multicultural London English?,” *University of Pennsylvania Working Papers in Linguistics*, vol. 17, no. 2, pp. 165–174, 2011.

[5] N. J. Young, “The sociolectal and stylistic variability of rhythm in Stockholm,” *Language and Speech*, Jan. 2021, doi: 10.1177/0023830920969727

[6] K. L. Pike, *The Intonation of American English*. Michigan: University Press, 1946.

[7] D. Abercrombie, *Elements of General Phonetics*. Chicago: Aldine, 1967.

[8] C. J. Hoequist, “Syllable duration in stress-, syllable- and mora-timed languages,” *Phonetica*, vol. 4, pp. 203–237, 1983.

[9] R. M. Dauer, “Stress-timing and syllable-timing reanalyzed,” *Journal of Phonetics*, vol. 11, no. 1, pp. 51–62, Jan. 1983.

[10] P. Bertinetto, “Reflection on the dichotomy ‘stress’ vs. ‘syllable-timing’,” *Revue de Phonétique Appliqué*, vols. 91–93, pp. 99–130, 1989.

[11] P. Auer, “Is a rhythm-based typology possible? A study of the role of prosody in phonological typology,” *KontRI Working Paper*, vol. 21, pp. 1–101, 1993.

[12] J. Caro Reina and J. Szczepaniak, “Introduction: Syllable and word languages,” in J. Caro Reina and R. Szczepaniak (eds.), *Syllable and Word Languages*. Berlin: de Gruyter, pp. 8–42, 2014.

[13] S. Schmid, “Syllable typology and the rhythm class hypothesis: Evidence from Italo-Romance dialects,” in J. Caro Reina and R. Szczepaniak (eds.), *Syllable and Word Languages.* Berlin: de Gruyter, pp. 421–454, 2014.

[14] F. Ramus, M. Nespor and J. Mehler, “Correlates of linguistic rhythm in the speech signal,” *Cognition*, vol. 73, pp. 265–292, 1999.

[15] V. Dellwo, “Rhythm and speech rate: A variation coefficient for deltaC,” in P. Karnowski and I. Szigeti (eds.), *Language and Language Processing: Proceedings of the 38th Linguistic Colloquium*. Frankfurt a. M.: Peter Lang, pp. 231–241, 2006.

[16] E. Grabe and E. L. Low, “Durational variability in speech and the rhythm class hypothesis,” *Laboratory Phonology*, vol. 7, pp. 515–546, 2002.

[17] I. Maddieson and S. Easterday, “The effects of phonological structure on the acoustic correlates of rhythm,” in *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, China, Aug. 2011, pp. 623–626.

[18] P. Mairano and A. Romano, “Rhythm metrics for 21 languages,” in *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, China, Aug. 2011, pp. 1318–1321.

[19] S. Frota and M. Vigário, “On the correlates of rhythmic distinctions: The European/Brazilian Portuguese case,” *Probus,* vol. 13, pp. 247–275, 2001.

[20] L. White and S. Mattys, “Calibrating rhythm: First language and second language studies,” *Journal of Phonetics,* vol. 35, pp. 501–522, 2007.

[21] V. Dellwo and S. Schmid, “Speaker-individual rhythmic characteristics in read speech of German-Italian bilinguals,” in A. Leemann, M.-J. Kolly, S. Schmid and V. Dellwo (eds.), *Trends in Phonetics and Phonology.* Bern: Peter Lang, pp. 349–362, 2015.

[22] A. Arvaniti, “The usefulness of metrics in the quantification of speech rhythm,” *Journal of Phonetics*, vol. 40, pp. 351–373, 2012.

[23] W. Labov, *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press, 1972.

[24] M.-A. Morand, M. Bruno, N. Julmi, S. Schwab and S. Schmid, “Speech rhythm in multiethnolectal Zurich German,” in *Proceedings of the 10th International Conference on Speech Prosody,* Tokyo, Japan, May 2020, pp. 566–570.

[25] M. Morand, S. Schwab and S. Schmid, *“*The perception of multi-ethnolectal Zurich German: A continuum rather than clear-cut categories,” *Loquens*, vol. 7, no. e072.

[26] D. Nübling and R. Schrambke, “Silben- versus akzentsprachliche Züge in germanischen Sprachen und im Alemannischen,” in E. Glaser, P. Ott and R. Schwarzenbach (eds.), *Alemannisch im Sprachvergleich*. Stuttgart: Franz Steiner Verlag, pp. 281–320, 2004.

[27] B. Siebenhaar, “Phonological and phonetic considerations for a classification of Swiss German dialects as a word language or a syllable language,” in J. Caro Reina and R. Szczepaniak (eds.), *Syllable and Word Languages*. Berlin: de Gruyter, pp. 327–345, 2014.

[28] S. Schmid, “Un nouveau fondement phonétique pour la typologie rhythmic des langues?,” Poster presented at the *10th Anniversary of the Laboratoire d'Analyse Informatique de la Parole*, Lausanne, Switzerland, May 2001.

[29] A. Leemann, V. Dellwo, M.-J. Kolly and S. Schmid, “Rhythmic variability in Swiss German dialects,” in *Proceedings of the 6th International Conference on Speech Prosody*, Shanghai, China, May 2012, pp. 607–610.

[30] A. Leemann, V. Dellwo, M.-J. Kolly and S. Schmid, “Disentangling sources of rhythmic variability between Swiss German dialects,” in *Proceedings of the 7th International Conference on Speech Prosody*, Dublin, Ireland, May 2014, pp. 693–697.

[31] U. Zihlmann, “Temporal variability in four Alemannic dialects and its influence on the respective varieties of Swiss Standard German,” in *Proceedings of the 10th International Conference on Speech Prosody*, Tokyo, Japan, May 2020, pp. 620–624.

[32] S. Schmid, “Obstruent voicing and temporal characteristics in Swiss German multiethnolects: A preliminary study,” Paper at *New Sounds 2013: International Symposium on the Acquisition of Second Language Speech*, Montreal, Canada, May 2013.

[33] H. Wiese and M. Pohle, “’Ich geh Kino’ oder ‘… ins Kino’?: Gebrauchsrestriktionen nichtkanonischer Lokalangaben,” *Zeitschrift für Sprachwissenschaft*, vol. 35, no. 2, pp. 171–216, Nov. 2016.

[34] M.-A. Morand, M. Bruno, S. Nadig, S. Schwab and S. Schmid, “Speech rhythm as a sociophonetic marker: A comparison of spontaneous and read speech in multiethnolectal Swiss German,” Poster presented at *Phonetics and Phonology in Europe (PaPE) 2021*, Barcelona, Spain, Jun. 2021.

[35] R. Baker and V. Hazan, “DiapixUK: Task materials for the elicitation of multiple spontaneous speech dialogs,” *Behavior Research Methods*, vol. 43, no. 3, pp. 761–770, 2011.

[36] K. J. van Engen, M. Baese-Berk, R. E. Baker, A. Choi, M. Kim and A. R. Bradlow, “The Wildcat corpus of native- and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles,” *Language and Speech*, vol. 53, pp. 510–540, 2010.

[37] K. L. Pike, “Etic and emic standpoints for the description of behavior,” in K. L. Pike (ed.), *Language in Relation to a Unified Theory of the Structure of Human Behavior*. Berlin, Boston: de Gruyter Mouton, pp. 37–72, 1967.

[38] P. Boersma and D. Weenink, *Praat: doing phonetics by computer* [Computer program], version 6.1.38, 2020.

[39] E. Dieth, *Schwyzertütschi Dialäktschrift: Dieth-Schreibung*, 2nd ed. by C. Schmid-Cadalbert. Aarau: Sauerländer, 1938/1986.

[40] T. Kisler, U. Reichel and F. Schiel, “Multilingual processing of speech via web services,” *Computer Speech & Language*, vol. 45, pp. 326–347, 2017.

[41] V. Dellwo, *Praat script: Duration Analyzer*, version 0.4, 2019.