SPATIAL VARIATION OF ARTICULATION RATE AND PHONETIC REDUCTION IN STANDARD-INTENDED GERMAN

Matthias Hahn and Beat Siebenhaar

Universität Leipzig, Germany
matthias.hahn@uni-leipzig.de, siebenhaar@uni-leipzig.de

ABSTRACT

Tempo of speech and phonetic reduction are closely related and differ in their spatial distributions. The *SpuRD*-project (*Sprechtempo und Reduktion im Deutschen*) focusses on this web of relationships and their spatial variation for the whole German-speaking area in central Europe. Using standard-intended reading material in normal and fast reading tempi, an array of reduction phenomena is analysed to find whether they are caused by tempo or whether they are independent variants of a limited linguistic area. The results reported here show on a macroscopic level that the spatial distributions of temporal characteristics such as the duration of articulation and the degree of segment reduction do not coincide everywhere, but have independent distribution areas especially at a higher tempo. That means that articulation rate is composed of regionally varying temporal and segmental features. For explanations of this macroscopic variation the material is analysed with regard to particular microscopic variation with independent spatial distributions.

Keywords: speech tempo, reduction, geophonetics, regional variation

1. INTRODUCTION

Speaking 'standard' means, at least for German, speaking without using salient regional pronunciation features [15, 11]. Most speakers are likely to use regionally accepted variants of the standard, some of which are well documented [1, 6]. However, regional prosody within the standard has not yet been investigated. There is much work concerning reduced forms, especially so called *weak forms* [14, 22]. However, these studies are mostly based on a corpus labelled as "standard material", which in fact contains mainly material from standard speakers of the northern German-speaking areas. Admittedly, there is a certain difference between "best standard in use" and its assessment [11]: the northern variant normally is assessed as the better standard. Nevertheless, variants of the middle and southern German area as well as those of Austria, Switzerland, Luxembourg, East Belgium, and South Tirol cannot simply be ignored. We therefore know little about segmental reduction behaviour in these regions and how it correlates to the codified or accepted reduction forms of the (northern) standard. We also know little about regional variation in temporal organization strategies, particularly the intrinsic tempo of spoken German, which is essential to temporal and segmental reduction processes. Finally, the entanglement of these factors when speech tempo is changed needs to be investigated.

2. THE MISSING LINK

2.1. Prosodic variation and space

The disadvantages of most of the studies considering speech tempo and space [17 for an overview] are twofold. Firstly, measurements of speech tempo are mostly set globally, meaning as average values by utterance. These measurements only find whether there are tempo differences, but not where these differences come from. Global measurements like articulation or speaking rates using words, syllables, or phones per time frame and including or excluding pauses have one thing in common: they do not reflect the details of varying temporal organisation strategies. For two hypothetical realisations of the same sentence from speakers of different regions we could obtain the same articulation rate value, but diverging temporal strategies. These could be, for instance, different final lengthening, varying intrinsic vowel or consonant durations, assimilations, deletions, or combinations of these factors. Therefore, we investigate both global statistical differences and the particular reasons for these differences.

Second, the conceptualisation of language space is often limited to the categories of traditional dialect classifications or even to countries or particular states, hindering insight into spatial continua. Moreover, traditional classifications of German dialect areas [2, 16, 21] are based on data collected in the late 19th century and are mainly based on phonological and morphological features. For this time span of about 140 years, one could likely assume some areal changes. What relationship do these traditional dialect areas have to prosodic variation in standard-intended German? It cannot be taken for granted that current variation in tempo-related features of standard-intended German necessarily coincide spatially with dated findings of phonological and morphological variation. Therefore, we use maps which are suitable for illustrating spatial continua and can be compared to the traditional dialect areas.

2.2. Combining the factors

Dressler [5] mentions that variants caused by fast speaking are often difficult to distinguish from seemingly identical variants of a particular regional dialect. He shows that Breton dialects exhibit regionally varying tendencies of vowel centralisation in fast speech ('allegro'). These findings on the relationship between area and reduction inspired us to identify regional variation in temporal pronunciation features in German-speaking Europe [10]. On the basis of 67 predominantly Federal German locations we can show (1) that there is regional variation for both articulation rate (segments/sec) and degree of reduction (measured as segment deletion ratio); and (2) that although these factors are closely interwoven, there is no simple correlation. The complex interplay between speech tempo, phonetic reduction, and regional variation needs to be studied in detail, especially in order to understand the composition of global measurements. This paper summarizes the progress of our investigations within the *SpuRD*-project.

3. SPURD-PROJECT

3.1. Aims of the project

The general aim of this project is to disentangle the interwoven relationships between speech tempo, reduction, and regional space, considering the whole German language area. Therefore, in a first step we test spatial variation for several measures of speech tempo and reduction indices on a macroscopic level. Since we can assume that this macroscopic variation accumulates from microscopic variation phenomena, we also study the regionally varying details of temporal organization as well as reduction preferences. Furthermore, we investigate how reading faster affects these spatial structures.

3.2. Material

We use reading material from the *German Today* corpus [4] which contains two separate readings of the German version of *The North Wind and the Sun,* which were read in a normal and a fast tempo. In total, 1,494 recordings are available from a younger (17–19 years, high school graduates) and an older generation (50–60, adult education classes), males as well as females, who originate from 195 locations in Austria, Belgium, Germany, Liechtenstein, Luxembourg, South Tirol, and Switzerland. However, the location density differs for the age groups, 165 locations are available for the younger generation and only 79 locations for the older generation. The subjects had no professional speaker training. The recordings were made in a formal interview situation during class in a quiet classroom. The preliminary results presented here are based on 274 already transcribed recordings of the younger males from 119 locations, which have been presegmented automatically at the sound level with WebMAUS [12] and manually corrected in Praat [3].

3.3. Methods: Measurement and mapping

Apart from the well-known disadvantages of using scripted and formal linguistic utterances, reading material provides the clear advantage of comparing variables one-to-one in the same environment. Geospatial comparability can only be achieved through this. The individual transcripts are compared with the canonical transcript that corresponds to the standard pronunciation. Thus it is possible to gain insights into spatial variation in the same linguistic elements, recognizing regional reduction strategies and not only average differences for a whole text or utterance. The accurate manual segmentation allows for measuring the sound segments in duration, pitch, intensity, and spectral qualities, as well as for global measurements on the text level for a general comparison.

Duration values for the sounds were automatically extracted with a Praat script. Segments that were labelled as slips of the tongue, as well as pauses that were not occlusions of plosives, were excluded. The maps presented in Figure 1 illustrate the spatial distributions of the following measurements: duration of articulation (*ArtDur* in sec) measures the time, minus pauses and slips of the tongue, that the speakers needed to read the text; articulation rate (*ARsyll* in σ/sec) measures the number of syllables realized per ArtDur; the reduction index segment deletion ratio (*SDR* in %) indicates the percentage of canonical sounds deleted.

We use ArcGIS [7] for mapping. The measured and averaged values for every location are transmitted to voronoi polygons to get a surface impression of the spatial distributions. A local smoothing method is further applied to flatten individual variation in the data. Hence, each polygon contains not only the values for the speakers of the location it represents, but the median value of itself and its direct neighbours. Finally, the data is classified into 10 classes with the *natural breaks* or *Jenks* algorithm [16], which is less arbitrary with class limits, reflecting instead variation within the data itself by minimizing the intra-class variation and maximizing the inter-class variation; this method helps to detect variance due to geographical space.

4. ANALYSIS

4.1. Macroscopic patterns

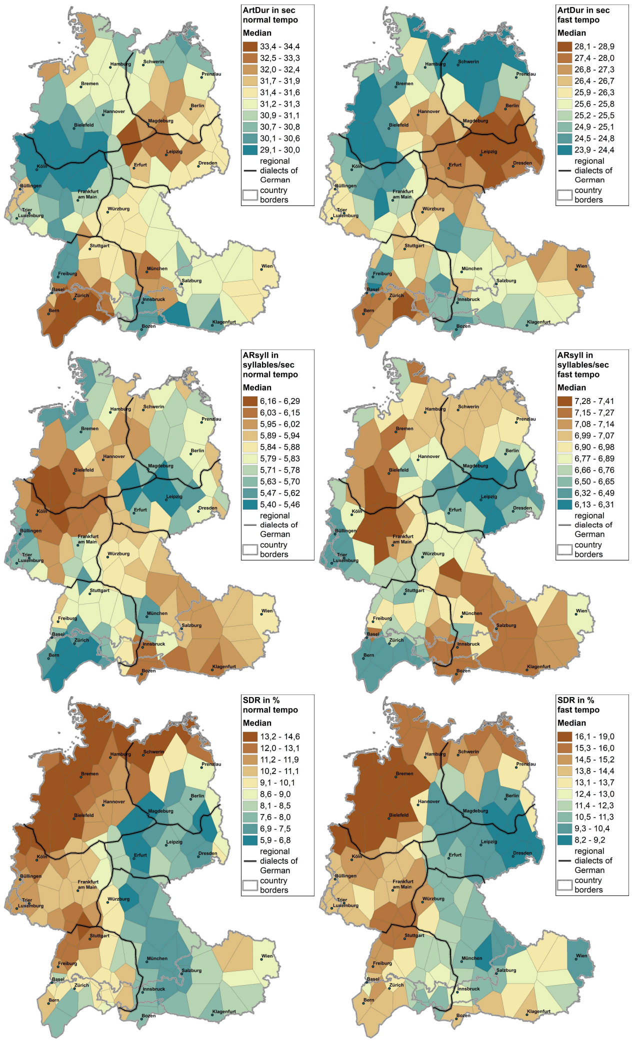
Table 1 gives an overview of the macroscopic measurements for the total data and subdivided for normal and fast tempo. As expected, the ArtDur decreases from normal to fast reading, while the values for SDR and ARsyll increase. In every case the ANOVAs show a highly significant influence of tempo. More interesting is the fact that the respective standard deviations increase with fast reading, indicating greater variation, which is a reinforcement of regional patterns for fast speech.

Table 1: ArtDur, SDR, and ARsyll, Mean and SD, ANOVA by tempo.

	ArtDur (sec)		SDR (%)		ARsyll (σ/sec)	
	Mean	SD	Mean	SD	Mean	SD
total	28.8	3.7	11.5	4	6.4	0.8
normal	31.5	2.2	9.9	3.2	5.8	0.4
fast	26.0	2.6	13.2	4.2	6.9	0.7
ANOVA by tempo	$F(1,273)$ =351.87, $p<.0001$*		$F(1,273)$ =53.26, $p<.0001$*		$F(1,273)$ =262.85, $p<.0001$*	

Figure 1 shows the spatial distributions for the global measurements of ArtDur (in sec, excl. pauses), the SDR (in %) and the ARsyll (in σ/sec, excl. pauses) for the normal and fast reading conditions. The classification of the values using the natural-breaks method makes it possible, despite significantly changed values between normal and fast reading, to map the variation inherent in the data. Figure 1 shows that the spatial structures for the measurement are highly congruent between tempos and sharper in shape when read fast, which partly explains the higher SD for the fast condition. This similarity of spatial structures strongly suggests that there is regional intrinsic variation with regard to temporal organization and the degree of reduction. Figure 1 also shows that the continua of value distributions do not correspond well with the black-marked borders between the traditional German regional dialects. This, in turn, suggests that prosodic variation in standard-intended pronunciation probably also shapes independent spatial structures that cannot be predicted from the traditional dialect areas. Table 2 lists the results of the ANOVAs by dialect area. The locations of the study area were assigned to the 6 large regional dialects (Low, Middle, and Upper German, each divided into an Eastern and Western part) according to traditional classifications [21]. For SDR, both ANOVAs yield highly significant results, while for ARsyll and ArtDur only those for normal tempo condition are significant. But considering the maps, the problem of this approach reveals itself. The dialect classification and the distributions of the mapped measurements rarely coincide. Rather, we find large variation within the dialect groups, especially for ARsyll and ArtDur. For SDR, the areas coincide better, i.e. with less internal variation, which in this case explains the high significance of the ANOVAs.

Figure 1: Overview: ArtDur, SDR, and ARsyll for the whole German language area in central Europe (on the left: normal tempo, on the right: fast tempo).



The spatial patterns for ArtDur and ARsyll match very well to each other, reflected in a strong negative correlation of r=-0.88 (Table 3). Areas in which speakers need more articulation time to read the text are simultaneously areas with low syllable rates, and vice-versa. However, such a correlation does not exist between the duration or tempo measures and the reduction index. Here, the correlations are regionally variable. For example, East Middle German shows high ArtDur (and low ARsyll), and according to the SDR values, low segmental reduction. East Upper German in the South, on the other hand, shows similarly low reduction values with on average significantly lower ArtDur (and higher ARsyll).

Table 2: One Way ANOVAs by dialect region for ArtDur, SDR, and ARsyll.

ANOVA	
ArtDur (normal)	$F(1,138)=3.28$, $p=.008$*
ArtDur (fast)	$F(1,134)=1.72$, $p=.1351$
ARsyll (normal)	$F(1,138)=2.64$, $p=.0261$*
ARsyll (fast)	$F(1,134)=1.04$, $p=.3996$
SDR (normal)	$F(1,138)=10.01$, $p<.0001$*
SDR (fast)	$F(1,134)=6.86$, $p<.0001$*

In addition, it is noticeable from Table 3 that the correlation strength decreases in all cases with higher speech rate, suggesting that temporal organization and segmental reduction are intertwined but nevertheless independent prosodic systems with independent spatial distributions.

Table 3: Pearson product-moment correlations.

Pearson product-moment correlations	
ArtDur : ARsyll (normal)	$r=-0.88$, $p<.0001$*
ArtDur : SDR (normal)	$r=-0.38$, $p<.0001$*
ARsyll : SDR (normal)	$r=0.23$, $p=.0068$*
ArtDur : ARsyll (fast)	$r=-0.70$, $p<.0001$*
ArtDur : SDR (fast)	$r=-0.34$, $p<.0001$*
ARsyll : SDR (fast)	$r=0.15$, $p=.0131$*

It is precisely this ‘dependent independence’ that makes this object of research so complex. Moreover, it is important to be aware that regions with macroscopically identical or similar values for articulation duration may have very different microstructures. The language material used here may not by any means capture all possible variation, but due to the direct comparability of the utterances it is appropriate to show reliable regional tendencies. In the following section, some of the microscopic variations are sketched, which in their sum lead to the macro-variation illustrated above.

4.2. Microscopic patterns: Exploring the reasons

In addition to the above analyses of global measures, different segmental, temporal, and spectral aspects are now analysed, allowing us to find partial explanations for geographical structures and detect interactions of different factors. So far in the SpuRD-project we have analysed the realisation of schwa between nasals [9]. The syncope of schwa is dependent on speech tempo, but there is also a geographical impact. While in the north schwa has disappeared in both tempi, it is preserved in the Bavarian area. In the transition zone, the syncope of schwa is more dependent on tempo, which means that schwa in these regions is more likely to become part of the reduction mass when the tempo is increased, while the Bavarian area remains largely insensitive to this form of reduction. Furthermore, we have studied the temporal realisation of the fortis-lenis contrast [8]. We find a geographical distribution of areas with both large and small temporal contrasts.

In both analyses, the spatial distribution is partially coherent with dialectal structures. Further analyses are planned, e.g. for the intrinsic duration of sounds and the realisation of the opposition between short and long vowels. We will try to identify those elements by factor analysis that can explain the main part of macroscopic variation. We also look at tempo-induced changes in sound quality. Elsewhere we show that vowel space size is geographically distributed – distinctly from traditional dialect boundaries – and that an increase in tempo results in a change of the vowel space which is dependent on the size of the vowel space in normal tempo, insofar as larger vowel spaces are reduced, while originally small spaces are enlarged [19].

5. CONCLUSION

Overall, this paper has shown that reading tempo, articulation rate, and segment reduction are clearly correlated, but that their linkage is complex, especially due to regional variability. Macroscopic and microscopic analyses must go hand-in-hand, as these relationships can only be disentangled with both perspectives. Therefore, in addition to investigating phonetics, the sociolinguistic dimension of this phenomenon must also be considered.

Furthermore, new methods, especially from the geosciences (such as hot spot analyses), should be tested and adapted for our purposes in order to be less dependent on both traditional and modern dialect classifications. These classifications are unsuitable for exploring the spatial structures of prosodic variation, as they are based on other linguistic levels and provide insufficient linguistic space for statistical analysis.

6. REFERENCES

[1] Ammon, U., Bickel, H., Lenz, A. N. 2016. Variantenwörterbuch des Deutschen. Die Standardsprache in Österreich, der Schweiz und Deutschland sowie Liechtenstein, Luxemburg,

Ostbelgien und Südtirol sowie Rumänien, Namibia und Mennonitensiedlungen. 2nd ed. Berlin: de Gruyter.
[2] Behaghel, O. 1891 Geschichte der deutschen Sprache. In: Hermann Paul (ed.), *Grundriß der germanischen Philologie*. vol. l, Straßburg: Trübner, 526–633.
[3] Boersma, P. and Weenink, D. 2017. Praat: doing phonetics by computer. Version 6.0.27. <http://www.praat.org/>
[4] Brinckmann, C., et al. 2008. German Today: an areally extensive corpus of spoken Standard German. *Proc. LREC* Marrakech.
[5] Dressler, W. U. 1975. Methodisches zu Allegro-Regeln. In: Dressler, W. U., Mareš, F. V. (eds.), *Phonologica 1972*, Salzburg: Wilhelm Fink, 219–235.
[6] Elspaß, S./Dürscheid, C. 2017. Areale Variation in den Gebrauchsstandards des Deutschen. In: Konopka, M., Wöllstein, A. (eds.): Grammatische Variation. Empirische Zugänge und theoretische Modellierung. Berlin/Boston: de Gruyter, 85–104
[7] ESRI 2015. ArcGIS Desktop: Release 10.3. Redlands, CA: Environmental Systems Research Institute.
[8] Hahn, M. in print. Zwischen Prozess und Produkt. Zur Lenisierung velarer Plosive im Deutschen. In: Hahn, M., et al. (eds.), *Dynamik in den deutschen Regionalsprachen.* Hildesheim: Olms.
[9] Hahn, M., Siebenhaar, B. in print. Schwa unbreakable – Reduktion von Schwa im Gebrauchsstandard und die Sonderposition des ostoberdeutschen Sprachraums. In: Kürschner, S., et al. (eds.), *Methodik moderner Dialektforschung.* Hildesheim: Olms.
[10] Hahn, M., Siebenhaar, B. 2016. Sprechtempo und Reduktion im Deutschen (SpuRD). In: Jokisch, O. (ed.), *Elektronische Sprachsignalverarbeitung 2016.* Dresden: TUDpress, 198–205.
[11] Kehrein, R. 2009. Dialektalität von Vorleseaussprache im diatopischen Vergleich – Hörerurteil und phonetische Messung. *Zeitschrift für Dialektologie und Linguistik* 76 (1). 14–54.
[12] Kisler, T., et al. 2017. Multilingual processing of speech via web services. *Computer Speech & Language* 45. 326–347.
[13] Kleiner, S. 2011ff. *Atlas zur Aussprache des deutschen Gebrauchsstandards (AADG).* http://prowiki.ids-mannheim.de/bin/view/AADG [Nov. 29, 2018].
[14] Kohler, K. J. 1995. *Einführung in die Phonetik des Deutschen*, 2nd edn. Berlin: Schmidt.
[15] Krech, E.-M., et al. 2009. *Deutsches Aussprachewörterbuch.* Berlin, New York: de Gruyter.
[16] Lameli, A. 2013 *Strukturen im Sprachraum. Analysen zur arealtypologischen Komplexität der Dialekte in Deutschland*. Berlin, Boston: de Gruyter.
[17] Leemann, A. 2016. Analyzing geospatial variation in articulation rate using crowdsourced speech data. *Journal of Linguistic Geography* 4 (02). 76–96.
[18] Lindblom, B. 1963. Spectrographic study of vowel reduction. *JASA* 35 (11), 1773–1781.
[19] Siebenhaar, B., Hahn, M. 2019. Vowel Space, Speech Rate and Language Space. In Calhoun, S., Escudero, P., Tabain, M., & Warren, P. (eds.), Proceedings of the 19th International Congress of Phonetic Sciences. Melbourne, Australia.
[20] Simmler, F. 1983. Konsonantenschwächung in den deutschen Dialekten. In: Besch, W., et al. (eds.), *Dialektologie: Ein Handbuch zur deutschen und allgemeinen Dialektforschung*, vol. 2, Berlin, New York: de Gruyter, 1121–1129.
[21] Wiesinger, P. 1983. Die Einteilung der deutschen Dialekte. In: Besch, W., et al. (eds.), *Dialektologie: Ein Handbuch zur deutschen und allgemeinen Dialektforschung*, vol. 2, Berlin, New York: de Gruyter, 807–900.
[22] Zimmerer, F. 2009. *Reduction in natural speech.* Dissertation, University of Frankfurt.