Vertical Differences in German Vowel Space Areas

Caroline Kleen[1], Marina Frank[1], Alfred Lameli[1]

[1]Research Center Deutscher Sprachatlas, Philipps-Universität Marburg, Germany

{caroline.kleen|marina.frank|lameli}@uni-marburg.de

Abstract

This study investigates phonetic variation in German dialects by analyzing vowel space areas (VSAs) across seven regional varieties and three communicative settings: dialect-intended speech (DIS), standard-intended speech (SIS), and read speech (RS). Results reveal significant vertical variation, with smaller VSAs in DIS compared to SIS and RS, reflecting reduced articulatory precision in informal registers. While vowel duration differs significantly across age groups, VSA configurations remain stable, indicating robust phonological structures. Regional variation in DIS highlights local phonetic distinctiveness, while overlaps in SIS and RS underscore the influence of linguistic standardization.

Index Terms: vowel spaces, vowel space area, regional variation, German

1. Introduction

In dialectology, a distinction is made between vertical and horizontal language variation [1]. Horizontal variation refers to differences across geographical regions, such as Low German dialects in northern Germany and Bavarian dialects in the southeast. Vertical variation, by contrast, encompasses the spectrum of language varieties between dialect and Standard German, reflecting influences such as linguistic proficiency, social identity, age, and the level of formality in communicative contexts [2]. This type of variation exhibits regional differences, with some areas displaying broader contrasts between dialects and Standard German than others.

The linguistic landscape in Germany has undergone substantial transformation over the past 130 years. Historically, spoken language was dominated by local dialects, while written language was accessible only to a small, literate segment of the population [3]. The introduction of standardized pronunciation in 1898 [4] and the spread of spoken Standard German through radio broadcasts since the 1930s [5] and television from the 1950s facilitated a shift toward bi-varietal language competence [6]. Today, speakers often navigate between dialects, Standard German, and intermediate varieties such as regiolects. These regiolects incorporate regional phonetic features, such as coronalization [7], and vary widely depending on the geographic area.

This study investigates vowel space areas (VSAs) to examine acoustic vowel variation across German regions. VSAs are defined as two-dimensional areas where vowels are represented by their first and second formant frequencies (F1 and F2). The F1 correlates with jaw openness (open–close dimension), while the F2 reflects tongue position (front–back dimension). Larger VSAs indicate greater use of the vocal tract and often correlate with higher intelligibility [8]. Beyond dialectological relevance, the analysis of vowel distributions contributes to the development of large language models (LLMs) and artificial intelligence (AI), which require robust training for dialect classification.

Previous research has documented VSA variation across languages and regions. Studies have identified dialect- and gender-specific VSAs, for example, in American English [9], regional differences in Australian English [10], and variability in VSA patterns in Peruvian Spanish [11], Maltese [10], and Portuguese varieties in Brazil and Europe [12]. Research on German and related varieties has revealed intergenerational changes in Luxembourgish vowel systems [13], VSA differences in Low German [14] and Bavarian [15]. These studies underscore the intricate interplay between region, register and phonetic variation. In [16], differences in vertical variation have been observed for speakers from Leipzig, depending on whether individuals employ standard-intended or dialect-intended speech. Furthermore, the VSAs for read speech were largest compared to the standard-intended and dialect-intended speech. As [17] demonstrates, also reading tempo can affect the vowel space differently across regions in Germany.

This study seeks to explore the effects of regional and contextual factors on vowel production. The hypothesis is that formant frequency values will exhibit both horizontal (regional) and vertical (register-based) variation, resulting in distinct VSA shapes and sizes.

2. Methods

2.1. Data

The data for this study come from the “Regionalsprache.de” (REDE) corpus [18], which contains audio recordings from 148 locations in Germany. The corpus includes three age groups: young speakers aged 18–23 (high school graduates), middle-aged speakers aged 45–55 (police officers), and older speakers aged 65+ (retired manual workers, such as farmers). For this analysis, gender differences cannot be examined as the corpus only contains male speakers. Participants are selected based on their origins, with priority given to individuals whose parents and grandparents are also from the same or nearby locations within the same dialectal region. Recordings are made using a Marantz professional PMD661 solid-state recorder. The specific settings can be found in [19].

In the present study, we analyze a subset of seven locations, chosen to represent different regions of the German language area as classified by [20]. These locations include:

10.21437/ISAPh.2024-9

- Low German: Northern Low German (location: Alt Duvenstedt), Mecklenburgish-West Pomeranian (Stralsund)
- Central German: Upper Saxon (Dresden), Moselle-Franconian (Wittlich)
- Upper German: East Franconian (Bamberg), Middle Bavarian (Trostberg), and High Alemannic (Waldshut)

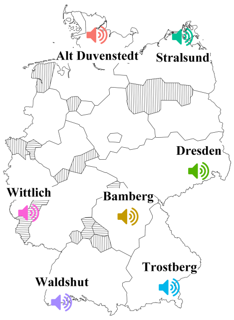


Figure 1: *REDE subset of seven locations in Germany, background: dialect classification by [20].*

We select two speakers per location, one middle-aged and one older speaker. The recordings capture three communicative settings: (1) Standard-Intended Speech (SIS): Participants listen to recordings of the Wenker sentences in their local dialect and translate them into their most formal Standard German; (2) Dialect-Intended Speech (DIS): Participants translate the Wenker sentences from Standard German, spoken by the interviewer, into their local dialect; (3) Read Speech (RS): Participants read aloud Aesop's fable The North Wind and the Sun [21]. We use the so-called Wenker sentences, developed by Georg Wenker in the late 19th century (see [22]), as a standard tool in German dialectology. These sentences enable us to systematically compare phonetic characteristics across registers and regions. By analyzing identical sentences in SIS and DIS, we investigate both vertical (register-based) and horizontal (regional) variation in vowel space areas.

2.2. Preprocessing and acoustic analysis

2.2.1. Formant extraction and vowel types

The dataset comprises 17,956 vowels, including 7,376 long vowels and 10,580 short vowels. To process this data, WebMAUS [23] was used for automated grapheme-to-phoneme conversion and forced alignment at both the word and phoneme levels. For the DIS data, orthographic transcriptions were based on dialect pronunciations (e.g., *flegen* instead of *fliegen* 'to fly'). Since WebMAUS was originally trained on standard language data, this required subsequent manual corrections to the resulting TextGrids. The final dataset included word and phone segmentation with SAMPA notation as corresponding labels.

To enable direct comparisons across SIS and DIS, each vowel was aligned with its Standard German reference vowel. This approach allows systematic analysis of identical items from the Wenker sentences across different registers. However, there might be allophonic shifts between dialectal and standard realizations. For example, the Standard German /ɪ/ may be realized as [ɛ] in DIS. These shifts affect the acoustic positions of allophones but do not impact the construction of vowel polygons, as all realizations of vowel categories are included. However, such shifts do influence the classification of vowel length, particularly when long monophthongs in the standard language are realized as short monophthongs in dialects. To address this issue, a phonetic classification into long and short vowels was performed. For 189 words (25% of the corpus), vowel length in the stem syllables was manually annotated for their dialect realizations, resulting in a total of 1,253 classified vowels (see Figure 2 for the manually classified data).

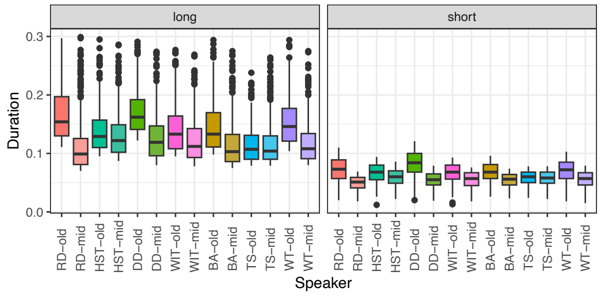
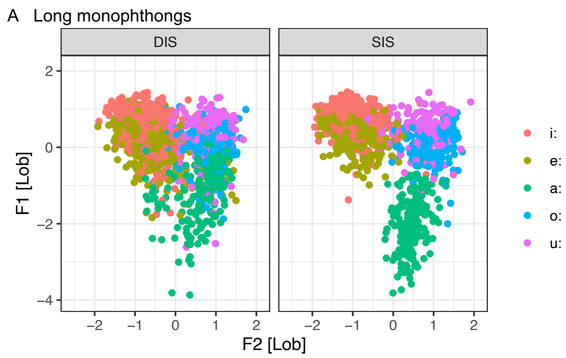


Figure 2: *Duration of long and short vowels per speaker.*

All analyses were carried out using the R programming environment [24]. Using this annotated subset, we trained a random forest classification model according to [25] in order to distinguish between long and short vowels for each speaker of the corpus. The model achieved 99.9% accuracy for the training data (70% of the manually pre-classified data) and 86.8% accuracy when speaker information was excluded. This procedure outperformed a threshold-based classification approach, which was based on the midpoint between the 25th percentile of long vowels and the 75th percentile of short vowels per speaker. The random forest model was subsequently applied to classify the remaining vowels in the corpus.

For the acoustic analysis, F1 and F2 values were automatically extracted using Praat [26] at 21 evenly spaced time points across each vowel. The midpoint values of F1 and F2 were used for further analysis. A formant ceiling of 5,000 Hz was set as the default for formant extraction, with adjustments to 4,500 Hz for the back vowels /oː, ɔ, uː, ʊ/ to improve accuracy. Figure 3 illustrate the distribution of the five corner vowels for long and short vowels in SIS and DIS. At first glance, the distribution of long vowels exhibits less variation than that of short vowels. Allophonic shifts in DIS, such as in the realizations of [iː] and [a], are evident but minor. These findings suggest that systematic phonetic differences between registers can be captured effectively through this analytical framework, providing insights into vowel dynamics across communicative settings.



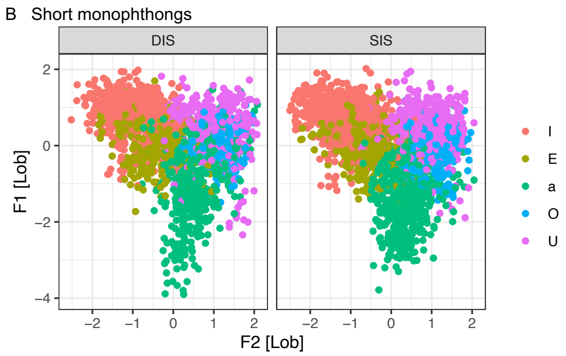


Figure 3: *Long and short vowels in dialect-intended speech (DIS) and standard-intended speech (SIS).*

2.2.2. *Identification of vowel space areas*

To ensure comparability across speakers, the F1 and F2 values of all vowels were normalized using the Lobanov method [27]. Lobanov normalization is particularly effective for cross-linguistic and cross-speaker comparisons [16], as it minimizes anatomical variation among individual speakers while preserving phonological and sociolinguistic differences in the data [28, 29]. To further improve data quality, extreme outliers for each vowel type were excluded based on the interquartile range (IQR) around their medians.

In order to calculate VSAs, we tested three techniques. (1) VSA-hull: Defined as the convex hull encompassing the mean values of all outermost vowels. (2) VSA-corner: Based on the convex hull of five corner vowels. For long monophthongs, the vowels /iː, eː, aː, oː, uː/ were used, while /ɪ, ɛ, a, ɔ, ʊ/ were used for short monophthongs (see Figure 3). (3) VSA-dist: Constructed using the Euclidean distances of the five vowels furthest from the centroid calculated by the mean F1 and F2 values for long and short vowels. The VSA was then calculated as the convex hull encompassing these vowels.

In total, 252 VSAs were calculated (3 communicative settings × 3 VSA methods × 2 vowel types × 14 speakers). A linear mixed-effects regression model was fitted to evaluate differences between the methods, with AREA as the dependent variable, METHOD as a fixed effect, and SPEAKER as a random intercept. We used the lme4 package for this [30]. Post hoc tests were conducted using the emmeans package [31]. As shown in Table 1, the statistical analysis revealed no significant differences between the three methods for either long or short monophthongs.

Given these results, VSA-corner was selected as the default method for subsequent analyses, as it requires fewer vowels while maintaining sufficient reliability for cross-register comparisons. This method also avoids potential complications associated with less stable vowel categories, such as /ɛː/, which have been noted in prior studies [15].

Table 1: *Pairwise VSA differences between VSA measurements (area ~ method + (1|speaker)).*

	β	**SE**	***p***
(Intercept)	2.287	0.082	<.001
VSA-corner – VSA-dist	-0.155	0.142	.518
VSA-corner – VSA-hull	-0.261	0.142	.159
VSA-dist – VSA-hull	-0.106	0.142	.737

3. Results

3.1. Age group

Table 2 presents the results of a linear mixed-effects regression model, which reveals a significant difference in vowel duration between age groups. In this model, DURATION is the dependent variable, AGE GROUP is the fixed effect, and SPEAKER is included as a random intercept. The β coefficient indicates a small overall effect, which becomes markedly stronger when short monophthongs ($t(7045.9) = -30.642$, $p < .001$) and long monophthongs ($t(10{,}059) = -49.707$, $p < .001$) are analyzed separately.

Table 2: *Pairwise VSA differences between age groups (duration ~ age group + (1|speaker)).*

	β	**SE**	**p**
(Intercept)	0.101	0.003	<.001
old – middle-aged	-0.017	0.004	<.001

However, this difference in duration does not manifest in the VSAs. Analysis using a linear mixed-effects regression model, with AREA as the dependent variable, AGE GROUP as the fixed effect, and SPEAKER as a random intercept, shows that VSAs are independent of age group. The results, summarized in Table 3, indicate no statistically significant differences between middle-aged and older speakers.

Table 3: *Pairwise VSA differences between age groups (area ~ age group + (1|speaker)).*

	β	**SE**	**p**
(Intercept)	2.799	0.082	<.001
old – middle-aged	-0.064	0.116	.590

3.2. Communicative setting

In contrast to age group, the communicative setting emerges as a significant predictor of VSAs. The analysis shows that DIS is distinctly different from the other settings, while SIS and RS exhibit a closer relationship. The results, summarized in Table 4, indicate that VSAs in DIS are significantly smaller compared to both SIS and RS. Additionally, there is a small but statistically significant difference between SIS and RS, with RS displaying slightly larger VSAs.

Table 4: *Pairwise VSA differences between communicative settings (area ~ situation + (1|speaker)).*

	β	**SE**	**p**
(Intercept)	1.903	0.076	<.001
DIS – RS	-1.518	0.096	<.001
DIS – SIS	-1.265	0.096	<.001
RS – SIS	0.253	0.096	.025

3.3. Vowel types

The analysis reveals significant differences in VSAs between long and short monophthongs (Table 5), with these differences varying across communicative settings (Figure 4). VSAs are consistently smaller in DIS compared to SIS and RS. Additionally, VSAs for short vowels are smaller than for long vowels in RS ($t(81.932) = -2.746$, $p = .007$) and SIS ($t(72.959)$

= -5.440, $p < .001$). However, this distinction is not statistically significant in DIS ($t(66.932) = -0.895$, $p = .374$).

Table 5: *Pairwise VSA differences between vowel types (area ~ vowel type + (1|speaker)).*

	β	SE	p
(Intercept)	2.175	0.084	<.001
long – short	0.265	0.077	<.001

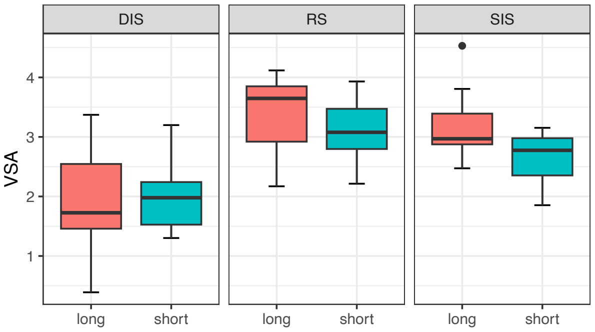


Figure 4: *VSA according to vowel type per communicative setting.*

The lack of significant differences between long and short vowels in DIS is unexpected and suggests the presence of phonological or articulatory simplifications in informal dialect speech. This contrasts with SIS and RS, where long vowels show more expanded VSAs, reflecting the increased articulatory effort associated with formal speech registers. Furthermore, the difference between short vowels in SIS and RS ($t(74.396) = 3.559$, $p = .001$) but not between long vowels ($t(81.995) = 0.999$, $p = .321$) may point to register-specific articulatory dynamics that affect vowel types differently. The individual polygons leading to this result are provided in Figure 5.

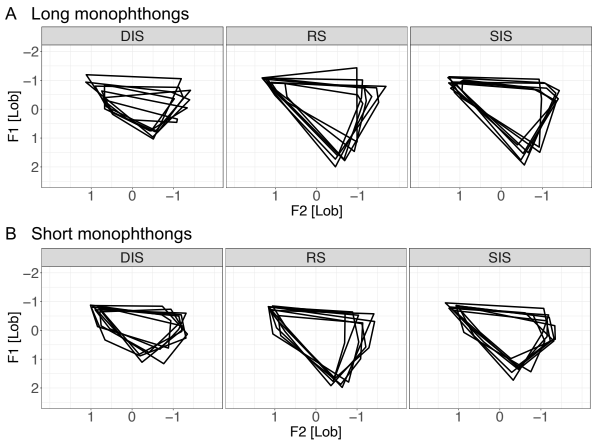


Figure 5: *VSA for vowels according to communicative settings of the seven locations (two speakers per location).*

4. Discussion

The finding that vowel space areas (VSAs) differ significantly across communicative settings highlights the strong influence of speech register on articulatory patterns. In dialect-intended speech (DIS), the reduced VSA may reflect an undershoot effect [32, 33], where speakers prioritize efficiency or naturalness over precise articulation. This contrasts with standard-intended speech (SIS) and read speech (RS), where larger VSAs suggest a more careful and deliberate articulation aimed at clarity. From a sociolinguistic perspective, the smaller VSA in DIS might also symbolize speakers' linguistic identity and adherence to regional norms. These findings have practical implications for technologies such as automatic speech recognition and synthesis, which must account for significant variation between registers to perform effectively in informal dialectal contexts. Future research could explore whether these patterns persist across a broader range of speakers and dialects.

The stability of vowel space configurations across age groups, despite notable differences in vowel duration, is a surprising outcome that challenges assumptions about intergenerational language change. The analysis by [34] demonstrates, however, that the same informants show significant differences between the age groups when the phonetic distance of vowel and consonantal realizations from the standard language is measured, rather than the VSA (see also [35]). The present result suggests that certain articulatory features, like VSA, may be deeply entrenched in linguistic systems, resistant to change even amid evolving speech habits or dialectal exposure. The lack of VSA variation may indicate a core phonological structure that remains consistent as part of speakers' fundamental language competence [36]. It contrasts with studies in other languages that have documented age-related vowel shifts [37], pointing to unique sociolinguistic dynamics in German dialects. Further research could examine whether this finding extends to regions with stronger generational divides in dialect usage, potentially revealing different dynamics in language stability and change.

Furthermore, the overlaps of the polygons should be pointed out (see Figure 5). The substantial overlap in VSA polygons observed in standard-oriented settings (SIS and RS) across regions indicates a high level of articulatory uniformity in formal contexts. This uniformity likely reflects the success of linguistic standardization efforts in Germany, supported by decades of formal education and national media promoting Standard German. In contrast, the greater variation in DIS reflects the persistence of regionally distinct phonetic features, underscoring the role of dialects in signaling local identity. The finding also raises questions about mutual intelligibility in these settings, as the overlap in formal registers may facilitate understanding across regions. This contrasts with the more localized variability seen in DIS, which could pose challenges for dialectal comprehension. Future studies could examine whether this overlap influences speakers' ability to adapt to unfamiliar dialects or impacts the perceptual boundaries between standard and dialectal speech.

The bi-varietal competence of the informants is also evident from an individual perspective. Figure 6 shows this by way of example for middle-aged speakers from the Upper German language area (Figure 6-A/B). The speakers from Trostberg (Central Bavarian) and Bamberg (East Franconian) are obviously able to switch between VSA configurations in DIS and SIS/RS. It becomes apparent that the standard and dialect systems in these regions are in a diglossic relationship to each other. This is in line with previous findings on the configuration of the variation spectrum between standard language and dialect [35]. For the data of the speaker from Dresden, however, it is not possible to clearly distinguish the communicative settings with regard to VSA (Central German, Figure 6-C).

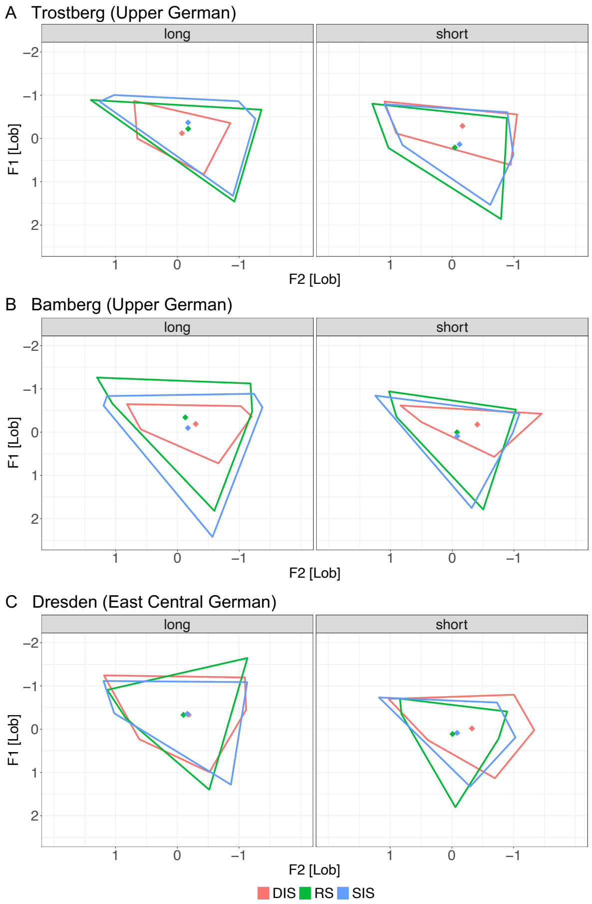


Figure 6: *VSAs for middle-aged individual speakers according to vowel types and communicative settings; centroids of VSAs as diamonds.*

This result corresponds to the VSA analysis by [16]. Furthermore, this is consistent with earlier work on the standard-dialect axis [34, 38], in which the East Central German language area shows only marginal language variation depending on the communicative setting, if at all. The stability of VSA in Dresden, across communicative settings, thus suggests unique sociolinguistic or phonetic dynamics in these areas. Figure 6 also indicates that the VSA positioning holds promise for deeper quantitative analysis. For example, while the VSA for long monophthongs in DIS in Dresden indicates a rather large jaw opening, the VSAs in Trostberg and Bamberg indicate smaller jaw opening within the same communicative setting.

Finally, the difference in VSA for short and long vowels should be highlighted. The observed VSA differences between long and short monophthongs align with established phonetic principles, as longer vowels typically exhibit more extreme articulatory targets (see [39] and [16]: 172 for an overview). However, the lack of a significant difference in DIS is unexpected and may point to particular phonological processes at play in informal speech contexts. Allophonic overlays in dialects might blur the distinction between vowel lengths, contributing to this finding. Alternatively, it could reflect a simplification in dialectal vowel systems, where phonological contrasts are less salient. These results highlight the need for further research into the phonological structures underlying dialects, particularly how vowel length interacts with other phonetic variables in shaping vowel space configurations. Exploring these dynamics could offer insights into how articulatory constraints and phonological systems evolve across communicative settings and regional varieties.

5. Conclusion

This study demonstrates significant vertical variation in vowel space areas (VSAs) across German registers, with smaller VSAs in dialect-intended speech (DIS) compared to the larger, more standardized VSAs in standard-intended speech (SIS) and read speech (RS). The findings underscore the influence of speech register on articulatory dynamics, with formal settings promoting greater articulatory precision and uniformity. Despite significant differences in vowel duration across age groups, VSAs remain stable, suggesting a robust phonological structure resistant to intergenerational change. The regional variation observed in DIS highlights the persistence of local phonetic characteristics, while the overlaps in SIS and RS reflect the homogenizing effects of linguistic standardization. These results provide valuable insights into the phonetic configurations of German dialects and their interaction with standard language, with implications for both sociolinguistic theory and practical applications in speech technology. Future research should expand the scope of analysis to include more regions, speakers, and demographic variables to deepen the understanding of phonetic variation in German.

6. Acknowledgements

This research is supported by the Federal Ministry of Education and Research (grant 16DKWN007 - Andy), the Academy of Sciences and Literature Mainz (grant REDE 0404), and the Research Center Deutscher Sprachatlas. We are grateful to two anonymous reviewers.

7. References

[1] A. Lameli, "Areale Variation im Deutschen „Horizontal“: Die Einteilung der arealen Varietäten des Deutschen," in *Sprache und Raum. Ein internationales Handbuch der Sprachvariation: Band 4: Deutsch*, J. Herrgen and J. E. Schmidt, Eds., Berlin, Boston: De Gruyter, 2019, pp. 185–205.

[2] R. Kehrein, "Areale Variation im Deutschen „Vertikal“," in *Sprache und Raum. Ein internationales Handbuch der Sprachvariation: Band 4: Deutsch*, J. Herrgen and J. E. Schmidt, Eds., Berlin, Boston: De Gruyter, 2019, pp. 121–158.

[3] R. Kehrein and H. Fischer, "Nähe, Distanz und Regionalsprache," in *Zur Karriere von ›Nähe und Distanz‹: Rezeption und Diskussion des Koch-Oesterreicher-Modells*, H. Feilke and M. Hennig, Eds., Berlin, Boston: De Gruyter, 2016, pp. 213–258.

[4] T. Siebs, *Deutsche Bühnenaussprache,* 11th ed. Bonn: Verlag Albert Ahn, 1915.

[5] J. E. Schmidt, "Vom traditionellen Dialekt zu den modernen deutschen Regionalsprachen," in *Vielfalt und Einheit der deutschen Sprache: Zweiter Bericht zur Lage der deutschen Sprache*, W. Klein, Ed., Tübingen: Stauffenburg Verlag, 2017, pp. 105–143.

[6] W. Besch and N. R. Wolf, *Geschichte der deutschen Sprache: Längsschnitte - Zeitstufen - linguistische Studien*. Berlin: Schmidt, 2009.

[7] J. Herrgen, *Koronalisierung und Hyperkorrektion: Das palatale Allophon des /ch/-Phonems und seine Variation im Westmitteldeutschen*. Stuttgart: Steiner, 1986.
[8] D. McCloy, R. Wright, and P. Souza, "Modeling intrinsic intelligibility variation: vowel-space size and structure," in *Proceedings of Meetings on Acoustics*, 2012, pp. 1–19.
[9] E. Jacewicz, R. A. Fox, and J. Salmons, "Vowel space areas across dialects and gender," in *16th International Congress of Phonetic Sciences, Saarbrücken, Germany*, 2007, pp. 1465–1468.
[10] D. Loakes and A. Gregory, "Acoustic Analysis of Vowels in Australian Aboriginal English spoken in Victoria," *Languages*, vol. 9, no. 9, pp. 1–29, 2024.
[11] E. O'Rourke, "Dialect differences and the bilingual vowel space in Peruvian Spanish," in *Selected proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology*, 2010, pp. 20–30.
[12] P. Escudero, P. Boersma, A. S. Rauber, and R. A. H. Bion, "A cross-dialect acoustic description of vowels: Brazilian and European Portuguese," *The Journal of the Acoustical Society of America*, vol. 126, no. 3, pp. 1379–1393, 2009.
[13] P. Gilles, "Ongoing sound changes in the Luxembourgish vowel system," in *Phonetik und Phonologie 17*, 2021.
[14] M. Gackstatter and O. Niebuhr, "Eine kontrastive phonetische Analyse niederdeutscher Langvokale," *Linguistik Online*, vol. 53, no. 3, 2012.
[15] S. Kleiner, "F1/F2-Diagramme als Darstellungsmittel bairisch geprägter standardsprachlicher Vokalsysteme," in *Germanistik*, Band 167, *Bayerisch-österreichische Varietäten zu Beginn des 21. Jahrhunderts – Dynamik, Struktur, Funktion: 12. Bayerisch-Österreichische Dialektologentagung*, A. N. Lenz, L. M. Breuer, T. Kallenborn, P. Ernst, M. M. Glauninger, and F. Patocka, Eds., Stuttgart: Franz Steiner Verlag, 2017, pp. 263–284.
[16] B. Siebenhaar, "Instrumentalphonetische Analysen zur Ausgestaltung des Sprechlagenspektrums in Leipzig," *Zeitschrift für Dialektologie und Linguistik*, pp. 151–190, 2014.
[17] B. Siebenhaar and M. Hahn, "Vowel Space, Speech Rate and Language Space," in *International Congress of Phonetic Sciences*, 2019, pp. 198–205.
[18] J. E. Schmidt, J. Herrgen, R. Kehrein, and A. Lameli, Eds., *Regionalsprache.de: Forschungsplattform zu den modernen Regionalsprachen des Deutschen*. Forschungszentrum Deutscher Sprachatlas Marburg, 2020ff.
[19] S. Lipfert, *REDE-Infothek: Dokumentation des Akademieprojekts Regionalsprache.de.* [Online]. Available: https://rede-infothek.dsa.info/
[20] A. Lameli, *Strukturen im Sprachraum: Analysen zur arealtypologischen Komplexität der Dialekte in Deutschland*. Berlin: De Gruyter, 2013.
[21] B. Ganswindt, R. Kehrein, and A. Lameli, "Regionalsprache.de (REDE)," in *Regionale Variation des Deutschen*, R. Kehrein, A. Lameli, and S. Rabanus, Eds., Berlin, Boston: De Gruyter, 2015, pp. 425–458.
[22] G. Wenker, *Schriften zum Sprachatlas des Deutschen Reichs*. Ed. by A. Lameli. Hildesheim, Zürich, New York: Olms (Deutsche Dialektgeographie), 2013.
[23] F. Schiel, "Automatic phonetic transcription of non-prompted speech," 1999.
[24] R Core Team, *R: A language and environment for statistical computing.* Vienna (Austria): R Foundation for Statistical Computing, 2021. [Online]. Available: https://www.r-project.org/
[25] A. Liaw and M. Wiener, "Classification and regression by randomForest," *R news*, vol. 2, no. 3, pp. 18–22, 2002.
[26] D. W. P. Boersma, *Praat: doing phonetics by computer [Computer program]*.
[27] B. M. Lobanov, "Classification of Russian Vowels Spoken by Different Speakers," *The Journal of the Acoustical Society of America*, vol. 49, 2B, pp. 606–608, 1971.
[28] C. G. Clopper, D. B. Pisoni, and K. de Jong, "Acoustic characteristics of the vowel systems of six regional varieties of American English," *The Journal of the Acoustical Society of America*, vol. 118, no. 3, pp. 1661–1676, 2005.
[29] M. Barlaz, *Vowel Normalization in R.* [Online]. Available: https://marissabarlaz.github.io/portfolio/vowelnormalization/
[30] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting Linear Mixed-Effects Models Using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.
[31] S. R. Searle, F. M. Speed, and G. A. Milliken, "Population Marginal Means in the Linear Model: An Alternative to Least Squares Means," *The American Statistician*, vol. 34, no. 4, pp. 216–221, 1980.
[32] K. N. Stevens and A. S. House, "Perturbation of vowel articulations by consonantal context: an acoustical study," *Journal of speech and hearing research*, vol. 6, pp. 111–128, 1963.
[33] B. Lindblom, "Spectrographic Study of Vowel Reduction," *The Journal of the Acoustical Society of America*, vol. 35, no. 11, pp. 1773–1781, 1963.
[34] A. Lameli, "Gesprochenes Deutsch in den Regionen: Eine Standortbestimmung für die Bundesrepublik Deutschland," in *Gesprochenes Deutsch: Struktur, Variation, Interaktion*, N. Proske, T. Weber, M. Dannerer, and A. Deppermann, Eds., Berlin, Boston: De Gruyter, 2025, pp. 51–79.
[35] R. Kehrein, *Regionalsprachliche Spektren im Raum: Zur linguistischen Struktur der Vertikale*. Stuttgart: Steiner, 2012.
[36] J. E. Schmidt and J. Herrgen, *Sprachdynamik: Eine Einführung in die moderne Regionalsprachenforschung*. Berlin: Schmidt, 2011.
[37] P. J. Watson and B. Munson, "A comparison of vowel acoustics between older and younger adults," in *Proceedings of the 16th International Congress of Phonetic Sciences*, 2007, pp. 561–564.
[38] M. J. Rocholl, *Ostmitteldeutsch - Eine moderne Regionalsprache?: Eine Untersuchung zu Konstanz und Wandel im thüringisch-obersächsischen Sprachraum*. Hildesheim, Zürich, New York: Olms, 2015.
[39] W. F. Sendlmeier and J. Seebode, "Formantkarten des deutschen Vokalsystems," *TU Berlin, Institut für Sprache und Kommunikation, Tech. Rep*, 2006.