Zeitschrift für
SPRACHVARIATION
& SOZIOLINGUISTIK
JOURNAL OF LANGUAGE
VARIATION AND SOCIOLINGUISTICS

https://doi.org/10.5282/jlvs/19

Phonetic Distance Outweighs Attitudes in Predicting Social Judgments of Swiss German Dialects

Tillmann Pistor[1], Adrian Leemann[1] and Beat Siebenhaar[2]

[1] Institute of German Studies, University of Bern, Bern, Switzerland
[2] Institute of German Studies, University of Leipzig, Leipzig, Germany
+
E-mail: tillmann.pistor@unibe.ch

Received 30.04.2026
Accepted for publication 05.10.2026
Published 05.11.2026

Abstract

Evaluations of regional speech variation emerge at the intersection of structural linguistic properties and socially shaped attitudes. While prior sociolinguistic research has emphasized the importance of social connotations, it remains unclear to what extent measurable phonetic differences themselves drive evaluative judgments. Against this backdrop, the present study investigates the role of (i) dynamic language attitudes toward dialect and (ii) segmental phonetic distance in predicting listeners' social evaluations of speech samples from two Swiss German varieties. Using an online perception task, we analyze positive and negative judgments by two listener groups from Switzerland and two from Germany (N = 61) evaluating matched-guise speech samples from Bern and Zurich dialect areas. We model the effects of individual regional backgrounds, attitudes and segmental phonetic distance on these evaluations in multiple regression and generalized linear models. Our results show that phonetic distance is the primary and region-independent predictor of evaluative judgments: greater distance reliably decreases positive ratings and increases negative ratings. It outweighs the influence of listeners' attitudinal predispositions and their familiarity with the respective dialects. These findings provide novel evidence for the central role of structural phonetic differences in evaluative perception and suggest a reassessment of the relationship between inherent values and social connotations in language evaluation.

Keywords: sociophonetics, language attitudes, phonetic distance

https://creativecommons.org/licenses/by/4.0/ Pistor et al.

1 Introduction

Swiss German dialects are regularly evaluated in public surveys and sociolinguistic research, often along aesthetic and socially transferred dimensions such as beautiful vs. ugly or friendly vs. arrogant (cf. e.g., Ris 1992; Berthele 2010; Christen 2010). Among these, Bern and Zurich German are particularly salient and frequently contrasted. This prominence reflects both socio-political visibility with Bern as the political capital and Zurich as the economic center, and socio-cultural stereotypes contrasting a rural and cozy Bern with a more rushed and urban Zurich. This perceptual prominence aligns with linguistic structure: both varieties occupy distinct

positions in Swiss German dialect space. Perceptual (Stoeckle & Schwarz 2019) and dialectometric research (Lameli et al. 2020) identifies them as belonging to separate clusters already at early stages of aggregation, reflecting well-known major west-east and north-south divisions (cf. Hotzenköcherle & Baumgartner 1962–1997; Christen 2019). These structural and perceptual differences are mirrored in subjective dialect rankings, in which Bern German is typically rated more favorably than Zurich German. However, such findings depend on methodology. Survey-based judgments relying on region labels often reflect reconstructed mental representations rather than responses to actual speech (cf. Hundt 2012), whereas evaluations based on authentic stimuli tend to be less negative and more favorable overall than those relying on reconstructions (cf. Purschke 2011). Sociophonetic research further demonstrates that listeners systematically link phonetic variation to social personae, shaped by stereotypes, context, and listener background (cf. e.g., Eckert 2008; Preston 2019; Sharma et al. 2022). For Swiss German, Pistor and Leemann (2024) show that specific phonetic features of Bern and Zurich varieties are associated with socially meaningful categories, while also emphasizing substantial individual variation in rating behavior.

Against this backdrop, a central question concerns the basis of evaluative judgments: are they primarily driven by perceived linguistic signal properties, or by pre-existing attitudes toward dialects and their speakers (cf. Purschke 2011, 2015)? Previous research typically frames this issue as an interaction between linguistic and social factors. In a famous study on aesthetic judgments of Dutch dialects, van Bezooijen (2002: 15) identifies phonetic form, contextual factors, cultural norms, intelligibility, and familiarity as key influences on the evaluation process.

Concerning phonetic form, the inherent value hypothesis (Giles et al. 1979; Giles & Niedzielski 1998) attributes evaluative effects to structural or acoustic properties of speech. Empirical support comes from phonaesthetic research showing cross-linguistic preferences for certain phonetic, phonotactic and sonority patterns (cf. Aryani et al. 2018; Reiterer et al. 2020; Kogan & Reiterer 2021, Anikin et al. 2023). In contrast, the social connotations hypothesis emphasizes extralinguistic associations, whereby evaluations rather reflect attitudes toward regions and speaker groups, crucially shaped by the evaluators' own regional origin (cf. Giles & Niedzielski 1998; Hundt 2010; Plewnia & Rothe 2012; Leemann et al. 2015). Closely related, the imposed norm hypothesis ("cultural norms" in van Bezooijen's list) posits that socially dominant or prestigious varieties are evaluated more positively due to their cultural status rather than linguistic properties (Giles et al. 1979: 406). Finally, familiarity with a given speech variety, is assumed to enhance positive evaluations (cf. van Bezooijen 2002: 15). In line with this assumption, recent research has shown that both the recognition of a language and the linguistic distance between the listener's L1 and the language being evaluated can predict evaluative ratings (cf. Reiterer et al. 2020; Mooshammer et al. 2023; Winkler et al. 2023).

Despite extensive international research, it is still unclear how strongly linguistic versus social factors contribute to evaluative judgments and how these components interact in the German-speaking context. In particular, the interaction between phonetic structure and language attitudes remains insufficiently explored. The present study addresses this gap by examining how phonetic distance as a linguistic factor and language attitudes as a social factor jointly shape evaluative ratings of Swiss German varieties. Focusing on Bern (BE) and Zurich (ZH) German, we analyze how phonetic (dis-)similarity, familiarity, and pre-existing attitudes toward regionally marked speech predict positive and negative evaluations. The study draws on four groups of native German speakers: two from Switzerland (Bern, Zurich) and two from Germany (Marburg/Hesse, Leipzig/Saxony).

The following sections first outline the role of attitudes in the respective speech communities and then operationalize familiarity as phonetic distance, forming the basis for the hypotheses tested in this study.

1.1 Language Attitudes and their sociolinguistic backgrounds in Switzerland and Germany

Across speech communities, speakers hold opinions about their own and others' language use, ranging from explicit to implicit forms depending on awareness and reflection (cf. Adler & Plewnia 2019: 141). These are commonly conceptualized as language attitudes (Eichinger et al. 2012), a "ubiquitous part of everyday

language use" (Pharao 2025: 350). Research over the past decades has established language attitudes as a core topic in sociolinguistics, closely linked to perceptual dialectology and folk linguistics (Preston 2019; Arendt 2019). A central distinction concerns attitudes toward entire varieties versus those toward individual linguistic features, the latter often discussed in terms of (social) indexicality, i.e., the "relationship between linguistic variation and socially meaningful categories" (Pharao 2025: 350). Earlier models conceptualize attitudes as relatively stable predispositions guiding evaluative responses (cf. Rosenberg & Hovland 1960: 1), whereas more recent approaches emphasize their dynamic and context-dependent nature. For instance, Purschke's REACT model (2015) treats attitudes as situationally constructed, while Soukup (2015, 2019) integrates both perspectives, still proposing cognitively grounded predispositions that however remain interactionally modifiable. A further conceptual distinction concerns the relationship between general language attitudes and the evaluation of specific linguistic stimuli, which differ in their attitude objects and in how decontextualized or context-bound the elicitation is (cf. Pharao & Kristiansen 2019). We will return to this distinction in section 2.3 after the description of which attitudinal data were gathered and how.

In the German-speaking area, evaluations of linguistic variation are closely tied to broader sociolinguistic dynamics (cf. Purschke 2020: 2). In what follows, we show that these dynamics differ substantially between Switzerland and Germany, shaping both attitudes and linguistic practices.

For German-speaking Switzerland, Steiner et al. (2023) demonstrate the impact of attitudes on linguistic practice, more precisely on the production of the indexically meaningful Swiss German variants /ç/ and /k/. Using a speaker-individual Dialect-Standard-Profile, they show that, alongside other social factors such as region, sex, and age, speakers' selection of the competing phonetic variants correlates with attitudinal scores: stronger dialect orientation predicts more dialectal realizations, while favorable attitudes toward the standard promote norm-oriented variants (cf. Steiner et al. 2023: 3545). This aligns with the well-described sociolinguistic situation in Switzerland, where dialect constitutes the primary variety of everyday communication. Its use is widespread, socially unmarked, and generally associated with high prestige (cf. Christen 2019: 246; Christen & Schmidlin 2019: 209; Bülow et al. 2021: 129). This situation largely contrasts with patterns observed in other parts of the DACH region, especially Germany (cf. Ammon 1995; Auer 2005).

Despite this overall positive valuation of dialect, internal differentiation exists. Between Bern and Zurich German, studies show that Bern German is generally viewed more positively, while Zurich German is perceived more negatively. While Bern German is typically associated with attributes such as beautiful, cozy, and friendly, Zurich German is described as brisk, harsh, or even arrogant (cf. Ris 1992; Berthele 2010; Christen 2010; Stoeckle & Schwarz 2019; Pistor & Leemann 2024). These patterns reflect the well-established status-solidarity distinction: Zurich German tends to score higher on status-related dimensions (e.g., competence), while Bern German is favored on solidarity-related dimensions (e.g., affective appeal; cf. Hengartner 1993). Yet again, these attitudes among Swiss German speakers vary systematically with raters' regional origin (cf. Stoeckle & Schwarz 2019).

In Germany, language attitudes are likewise shaped by social factors such as education, age, region, and metalinguistic knowledge (cf. e.g., Purschke 2011, 2015; Hundt et al. 2015; Hettler 2018; Kiesewalter 2019). However, the sociolinguistic configuration differs markedly from Switzerland. In Germany, use and active competence of dialects has declined substantially over the past 150 years, following a well-documented north-south gradient (cf. Schmidt 2017; Adler & Plewnia 2020). Consequently, the diglossic dialect-standard relationship observed in Switzerland does not apply to the Central German regions of Hesse and Saxony under investigation here.

For Hesse, Lameli (2025: 55) identifies diverse and highly variable repertoire types, reflecting considerable interindividual and situational variation. At the same time, dialect competence in Western Central German locations such as Hesse has declined sharply, especially among younger speakers (cf. Schmidt 2017: 117). The raters in the present study come from the Marburg area and thus from Central Hessian, a subregion of Western Central German. In his study of speakers from three generations in nearby Gießen, Vorberger (2020: 127)

finds that none of the speakers use dialect in everyday life and that everyday speech is predominantly regiolectal (a middle position within the vertical variational spectrum, cf. Kehrein 2015: 453–456) rather than dialectal.

By contrast, Saxony, located in the Eastern Central German area, shows a generally restricted repertoire spectrum (cf. Lameli 2025: 68). For Upper Saxon, under which the dialect of our participants' city of Leipzig is classified, the most prevalent variety has also been identified as regiolect (cf. Kehrein 2012: 213–233; Rocholl 2015: 309). Rocholl (2015: 309–310) shows that, irrespective of age, Eastern Central German speakers largely remain within this variety across communicative settings, while being able to shift stylistically within it, both toward and away from the standard. These findings are corroborated by Siebenhaar (2014) and Schaufuß (2015) for Leipzig specifically, as well as by Kehrein (2019: 127, 142). In sum, both German participant groups originate from regions in which dialect no longer plays a substantial role in everyday linguistic practice, and both are primarily familiar with the regiolect of their respective region.

This structural situation is reflected in evaluative patterns. In Germany, the standard variety generally carries the highest prestige (Adler & Plewnia 2018, 2019). Regional varieties, by contrast, are often evaluated less favorably. Saxon in particular is consistently rated lowest among German dialects and is associated with strongly negative attributes; Hessian varieties fare somewhat better but still attract substantial negative judgments (cf. Adler & Plewnia 2019: 145, 2020: 25). Especially Saxons have been confronted with negative attitudes toward their regional variety for well over a century and often tend to internalize these evaluations themselves (cf. Anders 2010a, b; Siebenhaar 2012: 160). It thus appears plausible that speakers' awareness of the broader reception of their variety may in turn shape their attitudes toward their own and possibly other dialects, even though it is not clear in which direction: Negative evaluations as in Saxony may be internalized and transferred to dialect as such, but may equally foster solidarity with other dialect speakers and more positive attitudes toward one's own dialect or dialect in general (cf. Anders 2010a). The specific outcome thus must be treated exploratorily.

With respect to German attitudes toward Swiss German, existing research points to very limited familiarity among Germans. Stoeckle (2014) shows that even German Alemannic speakers possess only vague and generalized knowledge of Swiss Alemannic dialects. The national border appears to function as a boundary of linguistic awareness, leading to superficial and unspecific representations of dialects beyond it. Similar findings are reported by Hundt (2010, 2017) and Schröder (2019). While Germans appear to possess a prominent cognitive representation of Swiss German as a whole, this representation appears to be largely generalized, which Hundt (2017: 128, translation TP) characterizes as an "unspecific intuition." Hundt (2010: 205–207) describes German representations of Swiss German as stereotypical but diffuse, combining both positive (e.g., smart, friendly, happy) and negative (e.g., unsympathetic, unpredictable, arrogant) attributes.

For Saxony, Siebenhaar (2012: 157) highlights additional comprehension difficulties: Swiss German is often not understood and may not even be recognized as German. Nevertheless, these perceptions coexist with generally positive attitudes toward Swiss people and, by extension, Swiss German speech.

This review of regional language attitudes provides the basis for hypothesizing that these attitudes, understood as cognitively grounded predispositions, may influence rating behavior. However, the direction and effect size of this influence cannot be reliably predicted a priori, as attitudes themselves are shaped by a wide range of linguistic and extralinguistic factors. The following section turns to linguistic factors that may serve as predictors of dialect evaluation.

1.2 Familiarity, similarity, and phonetic distance

As the goal of this study is to examine how phonetic similarity and dissimilarity predict evaluative behavior, this section outlines the theoretical and empirical background of the notion of linguistic "otherness" and its relation to familiarity (cf. van Bezooijen 2002; Mooshammer et al. 2023).

Familiarity as conceptualized by van Bezooijen (2002) as exposure to and knowledge of a speech variety, appears methodologically somewhat underspecified. It is not evident whether familiarity should be measured via self-report, whether it constitutes a binary or gradient

construct, listeners' ability to identify a given variety, or how "exposure" and "knowledge" are to be defined, ranging from frequency of contact to comprehension, recognition, or productive competence (cf. Hundt 2012: 202). A more precise and measurable construct closely related to familiarity is linguistic similarity, as both a perceivable and measurable quantity.

On the perceptual level, similarity is typically assessed using semantic differentials or comparable paradigms. In research on nonnative speech, it is often operationalized as perceived L2 accent strength, whereas in native speech contexts it reflects perceived differences between (national) standard varieties, regional accents, regiolects, and dialects within the same larger speech community, such as German in the DACH area.

A large body of research demonstrates that both segmental and prosodic phonetic variation plays a central role in social perception, categorization, and evaluation of speakers. This accounts for nonnative speakers' accents in the target language (cf. Fuertes et al. 2012; Wesolek et al. 2024) as well as for regionally marked speech of native speakers of the target variety (cf., e.g., Hettler 2018; Montgomery & Moore 2018; Sharma et al. 2022; Steiner et al. 2023). Linguistic cues can even outweigh visual information in social judgments (Rakić et al. 2011), and ingroup-outgroup categorizations operate rapidly, based on minimal acoustic input (Flege 1984; Park 2013). For Swiss German, Pistor and Leemann (2024) show that listeners assign social and aesthetic values to short stimuli within less than 1800 milliseconds. These findings highlight the reliability of linguistic cues in signaling social similarity and difference.

From an object-linguistic perspective, similarity can be quantified based on linguistic features. Dialectometric approaches, such as Lameli (2013) for German historical dialects model similarity using phonetic-phonological and morphological variables. Comparable methods have been applied to Swiss German (Scherrer & Stoeckle 2016; Lameli et al. 2020) and Austrian dialects (Vergeiner 2025).

Across these studies, phonetic-phonological variables consistently emerge as central components. Wieling et al. (2014), investigating dialectal sounds in data from six different language-corpora, combine both acoustically and segmentally determined sound differences with information-theoretic measurements to calculate phonetic similarity of the respective dialects. This exemplary study reflects a well-established method for operationalizing phonetic similarity as the calculation of phonetic distance between speech samples. In this type of measurement, two principal approaches can be distinguished: signal-based and segment-based methods.[1]

Signal-based approaches analyze acoustic properties (e.g., formant and fundamental frequencies) and have become increasingly visible in phonetic sciences across the globe (cf. e.g., Wieling et al. 2014; Bartelds et al. 2020; Lind-Combs et al. 2023; Clopper 2024). All newer studies employ dynamic time warping (DTW) to calculate distances based on entire speech signals, thereby incorporating phonemic, subphonemic, and prosodic information and is sometimes thus "referred to as holistic acoustic distance measure" (Lind-Combs et al. 2023: 4).

Segment-based approaches, by contrast, rely on symbolic representations of speech, most commonly using the Levenshtein algorithm (Levenshtein 1966), which compares sequences of phonetic segments and assigns costs to substitutions, insertions, and deletions.

Levenshtein-based measures can be implemented in unweighted or weighted form. While unweighted approaches assign equal costs to all operations, weighted variants differentiate between them based on linguistic relevance, often drawing on corpus-based evidence (Heeringa 2004; specifically for German varieties cf. Herrgen et al. 2001, Lameli 2004; Kehrein 2015; Levy et al. 2019). This allows researchers to account for the fact that certain phonological contrasts may be perceptually more salient or functionally relevant to listeners than others.

Both approaches face criticism from proponents of the alternative method. Segment-based methods are criticized for being prone to errors, labor-intensive and for neglecting prosodic information as well as acoustic detail (cf. Bartelds et al. 2020: 2), whereas signal-based approaches, particularly DTW, are sometimes considered insufficiently sensitive to linguistically meaningful distinctions (Lind-Combs et al. 2023: 3, 12). In a direct comparison of weighted and unweighted Levenshtein distances and DTW, Lind-Combs et al. (2023) found that weighted Levenshtein distance, as a segment-based approach, best captured listeners' perceived accent

distances for non-local native and nonnative accented American English. Lind-Combs et al. (2023: 10) themselves note, however, that the advantage of weighted over unweighted distances is small and may depend on language, task and stimulus type (e.g., sentences vs. words), and that the limitations of segmental and acoustic measures are partly complementary.

In German dialectology, weighted Levenshtein distance has been widely adopted since its introduction by Herrgen & Schmidt (1989) and subsequent refinements (Herrgen et al. 2001; Lameli 2004). Here, it has mostly been used to quantify dialectality as the distance between regional speech and the standard variety (cf. e.g., Kehrein 2012; Rocholl 2015; Schaufuß 2015; Vorberger 2019; Lameli 2025).

Beyond comparisons with the standard, phonetic distance can also be used to measure differences between regional varieties themselves. In an early study on interdialectal intelligibility in the Rhine- and Moselle-Franconian areas, Schmitt (1992: 103, 113–117) demonstrates that intelligibility decreases with increasing phonetic distance between varieties, particularly when differences involve systemic phonological contrasts. Kehrein (2015: 417) later confirms that strong phonological contrasts such as monophthongs vs. diphthongs or system-contrasting consonantal alternations play a central role in perception, whereas non-distinctive phonetic variation is less likely to affect comprehension and may primarily serve as a marker of regional origin. Kiesewalter (2019: 282–283) however showed that in perception experiments such weaker variants may nevertheless receive higher subjective dialectality ratings.

These findings support the use of weighted distance measures but also complicate predictions for social evaluation along a positive–negative scale. While phonetic distance has been linked to intelligibility, standard proximity, and accent strength, its role in shaping social and affective evaluations remains underexplored. One exception is Knops (1988), who shows for Dutch dialects that phonological deviations from the standard are evaluated more strongly than purely phonetic ones on status-related dimensions.

As the weighted Levenshtein approach forms the methodological basis of the present study, its implementation is detailed in Section 2.4. The choice of this approach also reflects the described research tradition and the comparability with existing data on German regional varieties it affords; beyond this, it is supported by empirical evidence: Importantly, these measures cross-regionally correlate with subjective perceptions of dialect strength (cf. Kehrein 2012: 330–332).

Overall, phonetic distance provides a theoretically grounded and empirically validated measure of linguistic difference. However, its role in evaluative perception and particularly its interaction with language attitudes has not yet been systematically investigated. This gap motivates the present study.

1.3 RQ and Hypotheses

The research question addressed in this study is whether and to what extent the two factors of (a) attitudes toward regionally marked speech and (b) measurable phonetic distance serve as predictors of positive and negative evaluations of Bern (BE) and Zurich (ZH) dialect speech samples. Based on the state of research outlined above, we formulate the following hypotheses:

- H1 (general predictors). Raters' region of origin, attitudes toward regionally marked speech and measurable phonetic distance function as predictors of evaluative rating behavior. This hypothesis follows from prior research suggesting that regionally differing attitudinal as well as purely linguistic factors may influence the perception and evaluation of speech varieties.
- H2 (attitudes and imposed norm hypothesis). Given the contrasting prestige configurations of the speech communities under investigation, overall attitudes toward dialect are expected to differ between rater groups. In German-speaking Switzerland, where dialect functions as the unmarked and socially dominant variety, predispositions as well as ratings are expected to be more positive than in Germany, where the standard variety occupies the primary prestige position.
- H3 (phonetic distance). Greater phonetic distance is expected to correlate with increasingly negative evaluations. This prediction is grounded in research showing that systemic phonological differences contribute more

strongly to perceived dissimilarity and social evaluation than non-systemic variation.

- H4 (phonetic distance, region and familiarity). The effect of phonetic distance is expected to be contingent upon region and, indirectly, familiarity. The German rater groups from Hesse and Saxony with lower presumed familiarity with Swiss German varieties than Swiss raters themselves are predicted to exhibit stronger distance-related effects.

2 Methods and Materials

2.1 Stimuli and participants

Following a matched-guise design (Lambert et al. 1960), stimuli were recorded as spoken by a professional actress. A total of 24 target words (e.g., *Tag* 'day', *Schwester* 'sister', *Kind* 'child'; for the complete list see Table 1 in Section 2.4) were recorded in both a Bern German and a Zurich German variant, yielding 48 stimuli in total. The lexical items were created such that each word contained at least one phonetic difference, which in most cases also is a phonological difference across the two dialects. We chose single words to explicitly construct matched minimal pairs across the two varieties, isolating the segmental features targeted by the phonetic distance measure detailed in Section 2.4. Consistent with standard matched-guise practice, this level of control also minimizes possible confounds in prosody, speech tempo, or lexical semantics, that operate above the segmental level and are not captured by that measure, while a fixed word list across speakers ensures replicability and direct comparability of stimuli.

As the stimuli consisted exclusively of content words, all items were controlled for potential inherent semantic affectivity using the Berlin Affective Word List (Võ et al. 2009) where applicable. All tokens were produced with a slightly rising intonation contour that closely resembles the word-list character and is usually interpreted as marking an enumeration. In disyllabic items, unmarked first-syllable stress was employed. All stimuli were controlled for accuracy and naturalness by L1 speakers of the respective dialects in the recording process.

A total of N=61 participants were included in the study, comprising both local and non-local native speakers of German. Approximately half of the sample was recruited in Switzerland, including participants from Bern (BE, n=19) and Zurich (ZH, n=12), while the remaining participants were recruited in Germany, including speakers from Hesse (HE, n=15) and Saxony (SA, n=15). 43 participants identified as female, 15 as male and 3 did not disclose their gender. Participants' age ranged from 16 to 36 years (M=25.79). Criteria for participant selection included comparable age, their own dialect being the one of the target region, and whether their parents spoke the same dialect as them. All participants had to be born, raised, and currently residing in their respective cities: Bern and Zurich for the Swiss groups, and Marburg (Hesse) and Leipzig (Saxony) for the German groups.

Participant selection was motivated by the markedly different sociolinguistic conditions of Swiss German and the Central German regions Hesse and Saxony, particularly regarding the status and use of regional varieties (see Section 1.1).

2.2 Collection of rating data

The present study employed an online perception experiment that replicates the design reported in Pistor and Leemann (2024). The task was implemented as an association paradigm inspired by the Implicit Association Test (Greenwald et al. 2003). Following a within-subjects design, all participants evaluated the (same) full set of 48 stimuli described in the previous section.

Before the experiment, participants were instructed to position their left and right middle fingers on the 'w' and 'p' keys, respectively. Prior to the main task, participants completed a categorization phase in which written items representing the later rating categories were classified binarily as 'good' or 'bad' (see Fig. 1, left) This step served to capture participants' valence-related associations with the rating categories and to enable a consistent binary distinction between positive and negative superordinate categories.

Following a brief training sequence, the association task was administered. Participants listened to one audio stimulus per screen and were to assign it to one of the categories displayed on the upper left (w) and upper right (p) screen corners as quickly as possible (see Fig. 1, right).

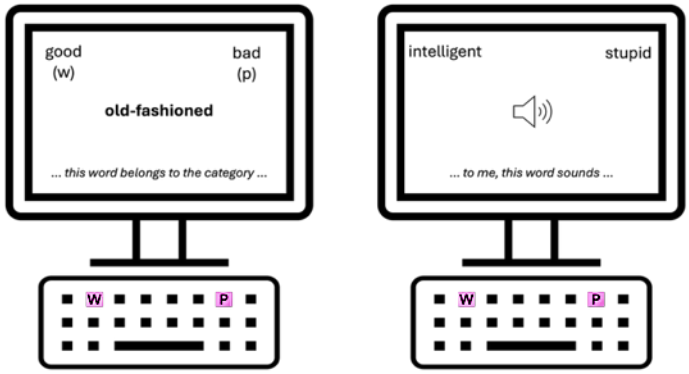


Figure 1: Experimental interface: left, binary valence classification of the rating categories; right, association task during stimulus presentation.

Six rating categories were included. Three represented aesthetic judgments commonly discussed in phonaesthetic research (beautiful/ugly, intelligent/stupid, attractive/unattractive, cf. Kogan & Reiterer 2021; Winkler et al. 2023). The remaining dimensions captured dialect-related attributes derived from dialectological and ethnological research (soft/hard, modern/old-fashioned, friendly/arrogant, cf. Ris 1992; Berthele 2010; Christen 2010; Stoeckle & Schwarz 2019).

Throughout the experiment, both category pairings and stimulus presentation were randomized for each participant. Participants' categorizations were recorded per click decisions. As general quality control measures, response times were monitored throughout the experiment to identify speeding or random responding and correlations between rating dimensions were examined to assess internal consistency.

2.3 Collection of attitudinal data

Following completion of the association test, all participants filled out a sociolinguistic background questionnaire. The attitudinal component comprised nine questionnaire items designed to capture respondents' orientations toward dialect in general and, more specifically, toward their own and other regional varieties (i.e. dialect affinity, cf. Hundt 2010).

In constructing this measure, we developed a dialect attitude profile based on an adapted version of the Dialect-Standard-Profile from Steiner et al. (2023, see Section 1.1). Substantial overlap exists between this instrument and the questionnaire items reported in Hundt (2010: 183–190), indicating that the measures draw on well-established constructs in language attitude research across the German-speaking area rather than being specific to Switzerland.

All items were rated on seven-point Likert scales. The questionnaire comprised the following items:

(1) I appreciate it when people from other regions recognize the dialect I speak.

(2) I appreciate it when people from other dialect regions speak their dialect with me.

(3) When speaking with people from another dialect region, I consciously avoid certain words or expressions.

(4) I consider my own dialect more aesthetically pleasing than others.

(5) I consider it regrettable when younger people are no longer familiar with many dialect words and expressions.

(6) Dialects naturally change over time; this does not concern me.

(7) I am concerned that dialect diversity in the German-speaking area may soon be lost.

(8) I would welcome greater uniformity if people across the German-speaking area spoke more similarly.

(9) I am proud to be a resident of my place of origin.

To enable numerical calculations, attitudes collected in the questionnaire were operationalized as a composite index. For this, participants' marks on the seven-point Likert scales in the nine questionnaire items were converted to numerical values and then summed to yield an overall attitude score, with a possible range from 9 (lowest possible absolute score) to 63 (highest possible absolute score). Items (3), (6), and (8) were reverse-coded, since agreement with these statements indexes a comparatively negative stance toward dialect and its preservation, whereas agreement with the remaining items reflects more positive attitudes. All absolute scores were subsequently normalized to a relative scale from 1 to 10. Higher attitude scores indicate a more positive orientation toward regionally marked speech.

In combination with rating data, we thus adopt two complementary perspectives on attitudes, most consistent with Soukup's (2019) integrative model. First, the questionnaire captures cognitively grounded predispositions toward unspecified dialects or regionally marked speech (memory-based attitudes). Second, the perception task elicits evaluations emerging from social categorization processes during direct exposure to speech stimuli (context-dependent attitudes).

Both measures elicit attitudes explicitly, but they differ in their attitude objects. The questionnaire targets dialects in general, one's own variety and those of others, and thus draws on a broad, decontextualized body of evaluative knowledge, whereas the perception task asks for judgments of specific stimuli on predefined binary categories, activating a narrower, context-bound body of knowledge (cf. Coupland & Bishop 2007; Pharao & Kristiansen 2019; Sharma et al. 2022). We thus argue that the two measures should therefore not be understood as identical manifestations of the same underlying process, even though they are nevertheless interconnected by the topic of regional speech variation. A separation of these two levels still appears methodologically and theoretically justified. In the analyses and discussion sections, we will examine how these two facets of attitudes interact.

2.4 Phonetic distance measurement

We applied the phonetic distance measurement procedure proposed by Herrgen and Schmidt (1989), Herrgen et al. (2001) and Lameli (2004). Detailed descriptions of the method can be found in the respective literature and in Kehrein (2012). The system defines class and level weights for vowels and dimension weights for consonants. For vowels, class refers to (a) the front-central-back dimension and (b) the rounding of the vowel within the IPA vowel trapezoid, while level denotes the vertical position in the trapezoid (close, close-mid, open-mid, open). For consonants, the relevant dimensions are place of articulation, manner of articulation, and ± voicing, as represented in the IPA chart of pulmonic consonants. Generally, differences across defined classes or levels are assigned a cost of 1.0, while diacritic-marked deviations (i.e., gradient tendencies) are assigned a cost of 0.5. The measure operationalizes segmental differences between transcribed realizations and does not model real-time auditory processing during lexical access. Its perceptual relevance is supported at the outcome level: Distances derived with this method predict decreasing intelligibility between varieties (cf. Schmitt 1992; Herrgen et al. 2001) and correlate strongly with listeners' judgments of dialectality (cf. Kehrein 2012: 330–332). The complete empirically motivated weighting scheme applied in the Levenshtein distance measurement is summarized in Table 1.

Table 1: Phonetic distance weighting scheme for vowels and consonants.

Monophthongs: Feature / Operation	**Cost**
Difference across vowel class/level (full step)	1.0
Difference across vowel class/level (half step; diacritic)	0.5
Quantity difference	1.0
Nasality difference (despite diacritic marking)	1.0
Full vowel vs. central vowel	1.0
Deletion of central vowel	1.0
Exceptions	
Centralized vowels (i/ɪ, y/ʏ, u/ʊ)	0.5
Near open ɛ/æ	0.5
Maximum cost constraints	
Maximum cost for /a/-differences	1.5
Maximum cost per vowel class/level dimension	1.5
Maximum cost per single monophthong segment total	3.0

Diphthongs: Feature / Operation	**Cost**
Individual component differences	= monophthongs
Monophthong vs. diphthong	3.0
Maximum cost constraints	
Maximum cost per single diphthong segment total	3.0

Consonants: Feature / Operation	**Cost**
Difference in place of articulation	1.0
Difference in manner of articulation	1.0
Difference in phonation (± voice)	1.0
Vocalized /r/ vs. consonantal /r/	1.0
Deletion/addition of homorganic affricate component	1.0
Diacritic difference	0.5
Consonant deletion/addition	2.0
Maximum cost constraints	
Maximum cost per single consonant segment total	2.0

For the measurement procedure, the original Bern (BE_real) and Zurich (ZH_real) stimuli were compared with realizations representing the Hessian and Saxon regiolectal varieties. For the latter two regions, additional recordings of four speakers per region (two male, two female) producing the same stimulus word list were collected and transcribed in IPA. Their segmental features were compared both internally and with descriptions reported in the dialectological literature (Schaufuß & Siebenhaar 2012; Rocholl 2015; Vorberger 2019; Ganswindt et al. 2023; Pistor et al. 2023). Where speakers diverged on a given segment, the variant

already attested in the literature for the respective age cohort was adopted. On this basis, representative HE and SA reference realizations ("dummy forms", HE_dum, SA_dum) were constructed. The dummy forms served as proxies for typical articulatory patterns of the respective varieties and enabled systematic phonetic distance calculations. They reflect contemporary realizations representative of our relatively young speaker cohort rather than historical dialect forms. While intra- and interindividual in natural speech variation cannot be captured by such abstractions, observed variation in the underlying transcriptions was minimal. As a corresponding simplification, phonetic distance within Swiss rater groups (e.g., BE listeners evaluating BE stimuli) was operationalized as zero, even though some residual variation would be expected in natural speech.

2.5 Statistical modeling

For the statistical analyses, phonetic distance was operationalized using the weighted Levenshtein distance values reported in Table 2. For each stimulus item, a region-specific distance value was calculated and assigned to all corresponding ratings from participants of that region, linking each response to the phonetic distance between the stimulus realization and the raters' regional variety. Attitudinal predispositions were operationalized as composite index values (relative attitude scores) described in Section 2.3.

The statistical workflow follows Lind-Combs et al. (2023), beginning with regression models assessing individual predictors and subsequently progressing to more comprehensive models. Data were analyzed using multiple linear regression (Pearson correlation-based) and generalized linear models (GLMs) as well as mixed models serving as robustness checks (GLMMs) with a Poisson distribution and log link function. Statistical significance was set at $\alpha=0.05$.

3 Results

3.1 Overall ratings and region

This section provides an initial overview of rating patterns across groups by examining the overall distribution of positive and negative evaluations. In total, N=17,575 utilizable ratings were collected (BE: n=5,472, ZH: n=3,464, HE: n=4,319, SA: n=4,320). Table 2 shows the relative proportions of positive and negative ratings by rater group.

A clear and expectable contrast emerges between Swiss and German raters. Both Swiss groups show pronounced positive majorities, whereas the Hessian group (HE) shows a majority of negative ratings. Ratings from the Saxon group (SA) are evenly balanced. One-sample proportion tests confirm that the within-group distributions differ significantly from chance for all groups except SA ($p<0.001$*** for BE, ZH, and HE; SA: $p=1$).

In addition to this, Figure 2 displays the relative distributions of positive and negative ratings by region of raters and the regional origin of the variants (stimuli) they evaluated.

Our results replicate the findings of earlier studies (Leemann et al. 2015; Stoeckle & Schwarz 2019) and in detail those of Pistor and Leemann (2024: 449): Both Swiss rater groups exhibit a strong regional bias in favor of their own dialect features, with the effect being more pronounced among Zurich raters than among Bern raters. This regional bias is attenuated in both German rater groups: Neither the Hessian nor the Saxonian raters show a preference for dialectal material from either Swiss region, although both groups tend to evaluate stimuli containing Zurich German variants slightly

Table 2: Distribution of relative positive and negative ratings by rater group.

rater group	n ratings	positive ratings	negative ratings
BE	5,472	59.87%	40.13%
ZH	3,464	63.53%	36.47%
HE	4,319	46.61%	53.39%
SA	4,320	50.00%	50.00%

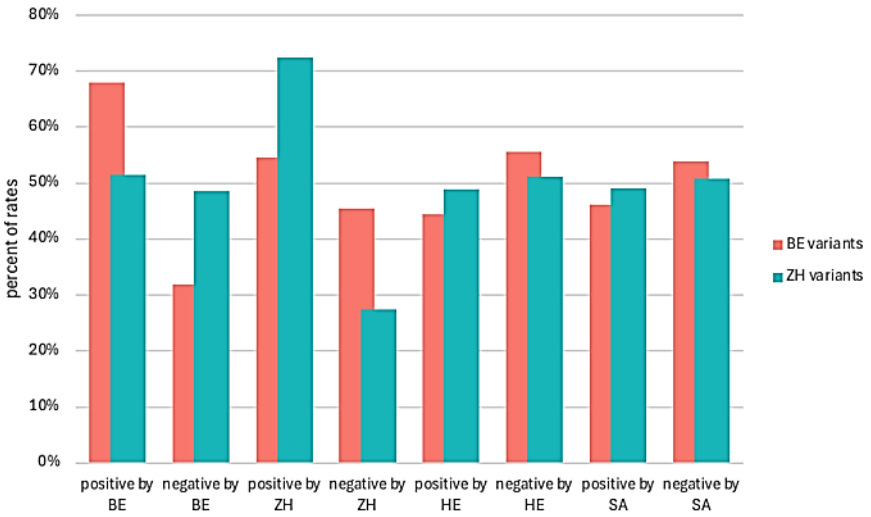


Figure 2: Distribution of relative positive and negative ratings by variants and rater group.

less negatively than those featuring Bern German variants. Thus, as expectable from previous research, the rating behavior observed within the Swiss rater groups is not reflected in the German rater groups.

The patterns observed here suggest that the raters' regional background is a relevant factor in the evaluation of Bern and Zurich German dialect samples. Swiss participants show a clear in-group preference, while German participants exhibit a more neutral pattern. These findings indicate that regional origin may serve as an important predictor for rating behavior and should therefore be taken into account in the subsequent analyses of attitudes and phonetic distance.

3.2 Attitudes and region

In this section, we first present the results from the attitude questionnaire, capturing participants' predispositions, and then examine how these attitudes split by region function as a predictor of positive and negative evaluations.

As expected, and formulated in hypothesis 2, overall attitudes toward dialect clearly differ between rater groups: participants from German-speaking Switzerland (BE and ZH) show a substantially more positive stance towards regionally marked speech than participants from Germany (HE and SA).

The boxplots in Figure 3 show the distribution of relative attitude scores across rater groups, clearly indicating systematic differences. Table 3 shows corresponding descriptive statistics for the overall attitude scores by rater group. While the means and medians of the two Swiss groups appear quite similar at first glance, analyses of variances indicate that the differences between all groups are highly significant ($p<0.001$***). Bern raters display the widest range of scores, with both the lowest minimum and the highest maximum values among all groups. In comparison, the standard deviation among Zurich raters is substantially lower than among Bern raters (ZH: SD=0.71, BE: SD=1.34), indicating that the Zurich group shows considerably less variation in their scoring. For the German groups, participants from Saxony showed a significantly higher mean attitude score and thus a more positive attitude toward regionally marked speech than those from Hesse.

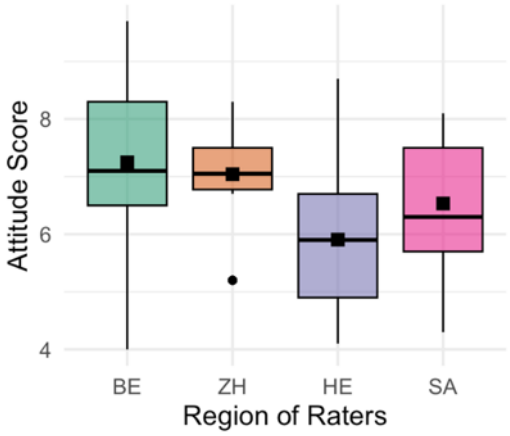


Figure 3: Distribution of relative attitude scores across rater groups.

Table 3: Descriptive statistics for relative attitude scores.

rater group	mean	median	min	max	sd
BE	7.25	7.10	4.00	9.70	1.34
ZH	7.04	7.05	5.20	8.30	0.708
HE	5.91	5.90	4.10	8.70	1.14
SA	6.53	6.30	4.30	8.10	1.05

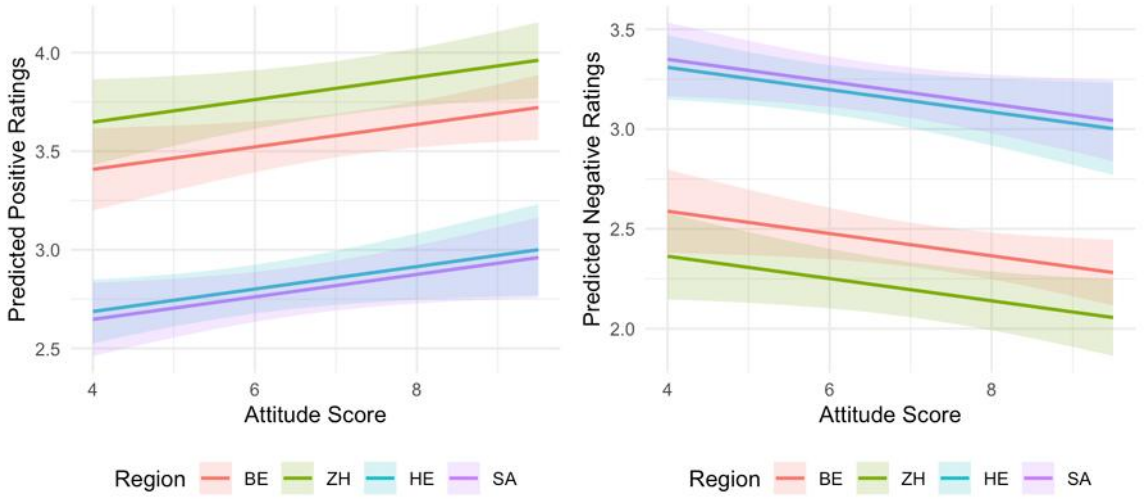


Figure 4: Regression plots for models predicting positive and negative ratings from attitude score and rater region.

To examine whether participants' attitude scores predict their ratings of positive and negative evaluations, linear regressions were fitted separately for positive and negative ratings (number of ratings per participant-stimulus combination, N=2,928 observations), with attitude score and raters' region (dummy-coded, reference: Bern) as predictors (see Fig 4 and Tab. 4). Because positive and negative ratings are complementary, the two models do not constitute independent evidence but mirror each other. Here and in our next models, both are reported to show that the pattern holds on either response scale. For positive ratings, attitude scores were a significant but modest predictor (β=0.057, p=0.041*), with higher attitude scores associated with slightly more positive ratings. For negative ratings, attitude scores were also significant (β=-0.056, p=0.045*), indicating vice versa that higher dialect affinity was linked to fewer negative evaluations.

However, the effect of rater region is substantially stronger in both models, showing that the regional

Table 4: Results from two regression models predicting positive and negative ratings from attitude score and rater region (reference category: BE).

	DV: positive ratings				DV: negative ratings			
Predictor	β	sd	t	p	β	sd	t	p
Intercept	3.180	0.210	15.157	<0.001***	2.811	0.210	13.407	<0.001***
Attitude sc.	0.057	0.028	2.042	0.041*	-0.056	0.028	-2.002	0.045*
Region: ZH	0.240	0.090	2.665	0.0077**	-0.226	0.090	-2.509	0.012*
Region: HE	-0.721	0.092	-7.834	<0.001***	0.721	0.092	7.840	<0.001***
Region: SA	-0.761	0.086	-8.805	<0.001***	0.762	0.086	8.819	<0.001***
Adjusted R^2:	0.0668				0.0657			

background of the raters clearly overshadows the predictive contribution of attitude scores. The individual behavior of the respective regions is consistent with the regional distribution of positive and negative ratings shown in section 3.1.

Overall, the models only explained roughly 6–7% of the variance in ratings (adjusted R^2=0.0668 for positive ratings, 0.0657 for negative ratings), suggesting that while attitudes toward dialect are a significant predictor and participants' region of origin exerts an even larger influence on evaluation behavior, other factors must become relevant.

3.3 Phonetic distance and region

We first show the results of the phonetic distance measurement procedure. All phonetic realizations of the word list targets and their corresponding phonetic distance values as results of the weighted measurement procedure reported in section 2.4 are shown in Table 5.

Overall, the mean phonetic distance between the two Swiss groups (BE-ZH: x̄=1.85) is, as expected, lower than the distances observed between the German and Swiss groups. Both German (dummy) groups exhibit greater mean distances to Bern German variants (HE-BE: x̄=3.06; SA-BE: x̄=3.60) than to Zurich German variants (HE-ZH: x̄=2.38; SA-ZH: x̄=3.33).

Within the German rater groups, the Saxon dummy shows consistently higher mean phonetic distances to both Swiss German varieties than the Hessian dummy.

To examine the predictive role of the phonetic distance values for positive and negative evaluations, the regression models introduced above (same dependent variables, unit of observation and coding of region) were recalculated with phonetic distance as independent variable (see Fig. 5 and Tab. 6).

Table 5: Phonetic realizations of the word list targets and their corresponding phonetic distance values.

Target	BE_real	ZH_real	HE_dum	SA_dum	BE-ZH	HE-BE	HE-ZH	SA-BE	SA-ZH
Bruecke	brʏkʰ	brʊkʰ	pʰʁʏkə	bʁʏgə	2	3.5	5.5	4	5.5
Dach	dɑχ	dɒχ	daχ	dɑχ	1.5	0.5	1.5	0.5	1.5
Dispo	dɪʃpo	dɪʃpo	dɪsbo	dĭsboː	0	2.5	2.5	2.5	2.5
Dreck	dræk͡χ	dræ̞k͡x	tʰʁɛkʰ	dʁɛ̆kʰ	1.5	4.5	5	4.5	4.5
fuenf	fyːf	fɔ͡ɪf	fʏnf	fʏnf	3	3.5	5	4	5
hart	hɛɹtʰ	hɛrtʰ	hɑːĕtʰ	hɑːʁˤt	1	4	4	5.5	5.5
Hund	hʊŋ	hʊntʰ	hʊntʰ	hʊntʰ	3	3	0	3	0
Kaese	xæːs	xɛːs	kʰɐːzə	kʰeːzə	0.5	4	3.5	4.5	4
Kasper	χɑʃβər	xɑʃbər	kɑsbɐ	kʰɑsbɐ	1	5	5	5	5
Kind	xɪŋ	xɪntʰ	kʰɪntʰ	kʰɪntʰ	3.5	4	1.5	4	1.5
Klasse	k͡χɬɑs	k͡xɬas	kʰlasə	kʰlazə	1.5	3	2.5	4.5	4
mal	mɑːʊ	mɑːl̃	mɑːl	mɑːl	4	3	1.5	3	1.5
Milch	mɪʊχ	mɪlx	mĭlç	mɪlʃ	3	4.5	1.5	3.5	1.5
nett	nætʰ	netʰ	nɛtʰ	nɛtʰ	1.5	0.5	1	0.5	1
Rest	ræʃt	reʃtʰ	ʁɛstʰ	ʁɛst	2	3	3	3	3
Riss	ris	ris	ʁĭs	ʁĭs	1	2.5	2.5	2.5	2.5
Schwester	ʃvɐʃdəɾ	ʃvœʃdɜɾ	ʃvɛstɐ	ʃvɜstɐ	1.5	3.5	4.5	5.5	6
sechs	sæχs	sæxs	sɛk͡s	zɛk͡s	0.5	1	1.5	2	2.5
Tag	tɑːkʰ	dɒːkʰ	tʰaːkʰ	tʰaːg	1.5	1.5	2.5	2	3.5
Tanne	dɑnːɜ	dɑnɜ	tʰanə	tʰɑnə	1	2.5	1.5	4.5	3.5
uns	ʔyːs	ʔʊ͡ɪs	ʔʊns	ʔʊns	3	5	5	5	5
wichtig	vɪχtɪkʰ	vɪxtĭkʰ	vɪçtĭkʰ	vɪçtɪc	1	1.5	1.5	3.5	3.5
Winter	vɪntər	vintər	vɪntʰɐ	vɪntʰɐ	1.5	2.5	3	4	4.5
zaehlen	t͡sɛːʊːɜ	t͡sɛlɜ	t͡sɛːlɲ	t͡sɛːlɲ	4.5	5.5	3	5.5	3
x̄ phonetic distance					1.85	3.06	2.83	3.60	3.33

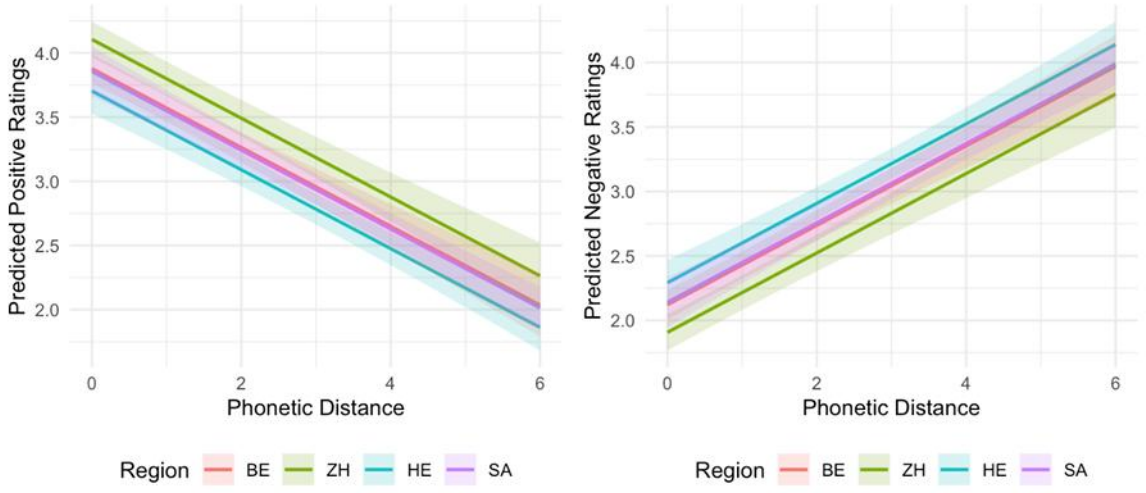


Figure 5: Regression plots for models predicting positive and negative ratings from attitude score and rater region.

The models reveal substantially stronger effects than those shown for attitudes. Phonetic distance is significantly associated with both positive and negative ratings and operates in the expected directions, as partially formulated in hypothesis 3: increasing phonetic distance increases negative evaluations (β=0.308, p<0.001***) and, as expected, decreases mirrored positive evaluations (β=−0.307, p<0.001***).

A further difference to the attitude models concerns the role of region of raters. Regional differences in direction and size of the effects are reduced once phonetic distance is included, and for some groups (HE, SA) they are no longer statistically significant. This suggests that part of the previously observed regional variation in evaluations is captured by phonetic distance itself. In other words, measurable phonetic differences between varieties appear to account for a substantial share of what initially appeared as mostly regional rating tendencies.

The reported differences are also reflected in the overall model fit. With about 12% (adjusted R^2=0.1222 and 0.1217), the phonetic distance models account for roughly twice as much of the variance than the attitude-based models. Phonetic distance thus emerges as a considerably stronger predictor of rating behavior than the attitudinal predispositions measured in the questionnaire: phonetic distance negatively predicts positive ratings and positively predicts negative ratings in all regions investigated.

3.4 Phonetic distance and attitudes

As a next step, generalized linear models (GLMs, Poisson distribution with log link) were fitted for positive and negative ratings as dependent variables, with phonetic distance and attitude score treated as fixed effects. Raters' region of origin was excluded to allow a more direct assessment of phonetic distance and

Table 6: Results from two regression models predicting positive and negative ratings from phonetic distance and rater region (reference category: BE).

	DV: positive ratings				DV: negative ratings			
Predictor	β	sd	t	p	β	sd	t	p
Intercept	3.878	0.058	66.880	<0.001***	2.121	0.058	36.616	<0.001***
Phon. dist.	-0.307	0.022	-13.753	<0.001***	0.308	0.022	13.802	<0.001***
Region: ZH	0.228	0.087	2.619	0.0089 **	-0.214	0.087	-2.461	0.0139 *
Region: HE	-0.174	0.093	-1.862	0.0627	0.171	0.093	1.830	0.0673
Region: SA	-0.021	0.099	-0.214	0.8309	0.019	0.099	0.192	0.8477
Adjusted R^2:	0.1222				0.1217			

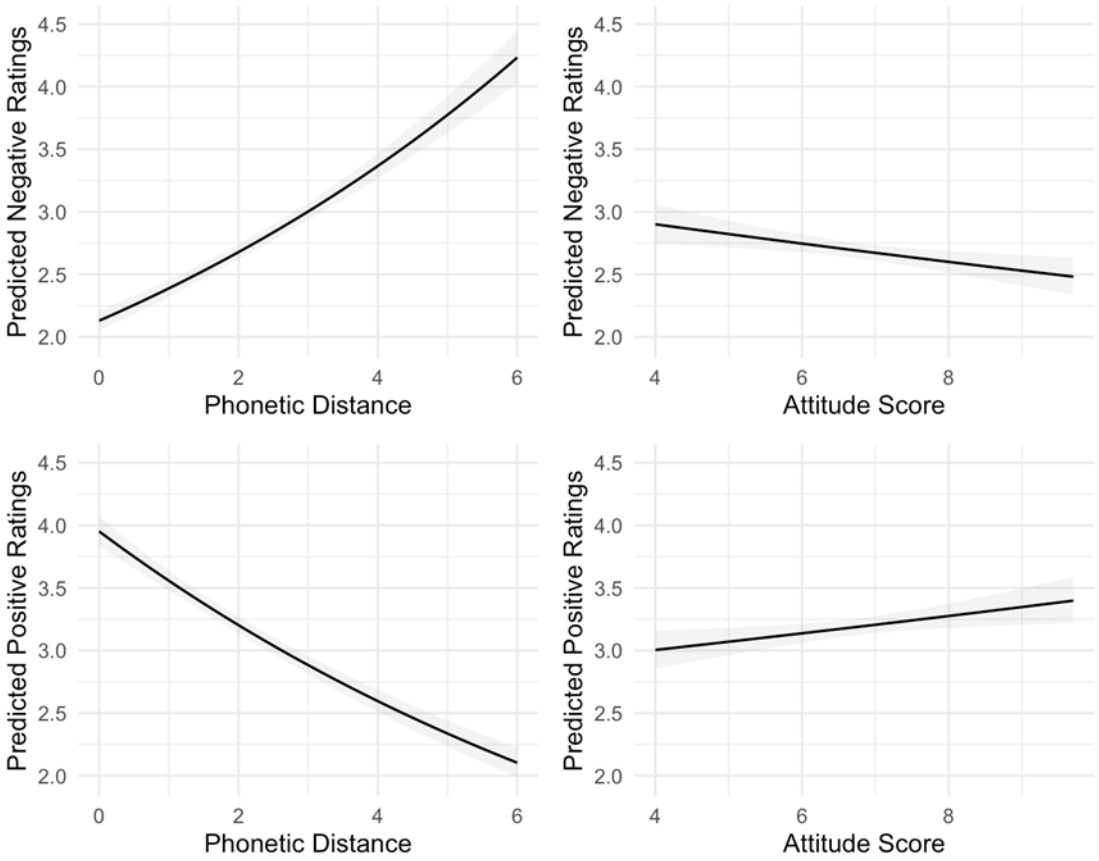


Figure 6: Plots for GLMs predicting positive and negative ratings from phonetic distance and attitude score.

attitudes. Although regional bias is evident (ingroup favoritism, see section 3.1), it is compensated in the operationalization of phonetic distance =0 for ingroup evaluations and primarily affects intercepts rather than effect estimates (see Fig. 5 and Tab. 6). Including region would thus mainly capture baseline differences without altering the effects of interest. To facilitate interpretation of the coefficients, percentage changes in the expected count were calculated as (eβ−1)×100, where β denotes the regression coefficient of the GLM. In subsequent models, interaction effects between the independent variables were additionally tested and checked for robustness in GLMMs.

In the fixed-effects generalized linear models (see Fig. 6 and Tab. 7), both phonetic distance and attitudes as independent variables exhibit significant effects. Without region as a factor, phonetic distance emerges as the strongest predictor in both models and shows expected opposing effects contingent upon the dependent variable. Phonetic distance has a significant negative effect for positive ratings (β=-0.105, p<0.001***), corresponding to an expected decrease of approximately 10.0% in positive ratings for each unit increase in phonetic distance. Conversely, for negative ratings, phonetic distance shows a significant positive effect

Table 7: Results from two GLMs predicting positive and negative ratings from phonetic distance and attitude score.

	DV: positive ratings				DV: negative ratings			
Predict.	β	sd	z	p	β	sd	z	p
Intercept	1.229	0.0621	19.79	<0.001***	0.939	0.0671	13.988	<0.001***
Phon. dist.	-0.105	0.0062	-16.93	<0.001***	0.114	0.0062	18.307	<0.001***
Attitude sc.	0.022	0.0085	2.53	0.0114 *	-0.027	0.0093	-2.948	0.00319 **

(β=0.114, p<0.001***), corresponding to an expected increase of about 12.1% per unit.

Attitude score also reaches statistical significance in both models, but its effect size is markedly smaller. Higher attitude scores slightly increase the likelihood of positive ratings (β=0.022, p=0.011*; ≈+2.2% per unit) and decrease the likelihood of negative ratings (β=-0.0273, p=0.003**; ≈-2.7% per unit). Thus, while attitudinal predispositions have a measurable influence on evaluation behavior, their contribution is substantially weaker than that of phonetic distance.

Overall, the results reinforce the earlier regression findings: phonetic distance constitutes the dominant predictor of evaluative behavior, whereas attitudes contribute only modestly to the prediction of positive or negative judgments in our test environment.

Next, the interaction between phonetic distance and attitude score was tested for both positive and negative ratings using GLMs. In both models, the interaction term (phonetic distance * attitude score) is non-significant (positive ratings: β=0.005, p=0.325; negative ratings: β=0.003, p=0.514), indicating no evidence that, in our sample, the effect of phonetic distance on ratings depends on participants' attitudes.

Finally, as a robustness check to account for the non-independence of repeated observations, both main-effects models were re-fit as generalized linear mixed models (GLMMs) with crossed random intercepts for participants and stimuli. The effect of phonetic distance remained significant and directionally consistent (positive: β=-0.093, p<0.001***; negative: β=0.099, p<0.001***), closely matching the fixed-effects estimates. Attitude score, by contrast, no longer reached significance in either model (positive: p=0.129; negative: p=0.243). This is expected, as each participant contributes only one attitude score: the effect rests on 61 individuals rather than on the 2,928 aggregated observations and is therefore estimated much less precisely. The mixed models thus sharpen the picture of the fixed-effects analysis: phonetic distance is a robust predictor of evaluative judgments, whereas the contribution of attitudes is small and not reliably established.

4 Discussion

In this final section, the four hypotheses are revisited to assess the extent to which they are supported by the present findings, followed by a consideration of the study's limitations.

H1 (general predictors) proposed that raters' region of origin, their attitudes toward regionally marked speech, and the measurable phonetic distance between varieties function as predictors of evaluative rating behavior. The results largely support this assumption. All three factors show significant effects on positive and negative evaluative judgments, indicating that, consistent with prior research (see section 1), both social and linguistic factors contribute to how Swiss dialect samples from Bern and Zurich are assessed. However, the effect of raters' region of origin is not uniform across participant groups. In the present data, this factor is primarily driven by the Swiss German participants: while German raters show little differentiation between Bern and Zurich German variants, Swiss raters clearly distinguish between them in their evaluations. This suggests that regional affiliation plays a stronger role among listeners who are themselves embedded in the dialect landscape in question. This asymmetry will be discussed in more detail below.

Building on these general predictors of rating behavior, the second hypothesis H2 (attitudes and imposed norm hypothesis) assumed that overall attitudes toward dialect would differ between rater groups due to the contrasting prestige configurations of the two speech communities. While dialect is the dominant variety in German-speaking Switzerland but not in Germany, attitudes and ratings were expected to be more positive among Swiss participants. Results clearly support this hypothesis: both Swiss rater groups show significantly higher mean attitude scores and thus more favorable predispositions toward regionally marked speech than the German groups.

One explanation for this general pattern may lie in the differing sociolinguistic situations of the regions involved. In the Central German areas of Hesse and Saxony, what is locally perceived as dialect often corresponds to regiolect situated along a dialect-standard continuum, rather than to a traditional base dialect. Here, regiolect can carry elements of hidden prestige, indexing solidarity and regional identity (cf. Anders 2010a: 267; Schaufuß & Siebenhaar 2012; Vorberger 2019), even though it lacks the broader social dominance of dialect found in German-speaking Switzerland.

In contrast, the Swiss context seems still to be more accurately described as a diglossic situation, in which dialect not only functions as the unmarked variety of everyday interaction but also carries aspects of overt prestige (cf. Christen 2019: 246). Against this background, the more positive attitudes towards dialect in general observed among Swiss participants can be understood as reflecting both the structural embedding of dialect in daily communication and its dual prestige status, whereas in the German context, positive orientations toward regionally marked speech are more closely tied to identity-marking functions within a less dominant, continuum-based system. At the same time, differences emerge within the Swiss sample. Bern raters display the highest mean attitude score, but also the largest internal variation, reflected in both the lowest minimum and the highest maximum score among all groups. Zurich raters, by contrast, exhibit slightly lower mean scores but considerably less variation, indicating a more homogeneous attitudinal profile.

Likewise, an interesting internal difference emerges between the two German groups. Saxon participants display more positive attitudes toward regionally marked speech than Hessian participants. This finding is noteworthy given that Saxon varieties are widely perceived as among the least popular dialects in Germany (cf. Adler & Plewnia 2019: 145). As Siebenhaar (2012: 160) stated, speakers of these varieties have been confronted with long-lasting negative evaluations of their speech and are often assumed to internalize such attitudes. In our present data, however, the pattern points in the opposite direction: rather than reinforcing negative attitudes toward dialect in general, this experience may foster comparatively positive predispositions toward regionally marked speech. One possible interpretation is that these attitudes function as a form of counter-positioning against the broader negative perception of the variety. This is a perspective that is hinted at in Anders (2010a: 267). Currently, a shift in the public perception is also being promoted through an official social media campaign by the Saxon state government titled "So geht sächsisch (simply Saxony)." At the same time, our sample consists exclusively of relatively young participants (19–27 years among the Saxon group), raising the possibility that generational change may play a role. Future research could examine whether similar patterns emerge across different age cohorts and could also incorporate attitudes toward the German standard variety for comparison.

Finally, the relationship between predispositions and actual rating behavior appears broadly consistent but not perfectly aligned. Overall, the models show that higher attitude scores indicating a more positive predisposition toward regionally marked speech predict more positive and fewer negative ratings, although this effect is overshadowed by raters' region of origin. Although attitude scores have a similar predictive effect across rater groups, differences in their baseline predispositions (visible in the intercepts) dominate the model.

The two Swiss groups, which show the most positive attitudinal predispositions, produce significantly more positive than negative ratings both within and across groups. Hessian raters, who exhibit less positive predispositions overall, provide fewer positive ratings than the Swiss groups and also produce more negative than positive evaluations within their group. Saxon raters show a perfectly balanced internal pattern, producing equal proportions of positive and negative ratings, while still assigning fewer positive evaluations than the Swiss groups. These tendencies suggest that self-reported predispositions are reflected in rating behavior to a considerable extent.

Furthermore, as noted before, our data shows a pronounced regional bias in ratings. Both Swiss groups evaluate their own dialectal variants more positively than those from the other Swiss region, whereas the German rater groups do not show such differentiation. This finding is consistent with previous research (e.g., Stoeckle & Schwarz 2019; Pistor & Leemann 2024) and suggests that Swiss listeners evaluate dialect samples with a finer-grained awareness of regional distinctions. German listeners, by contrast, appear to evaluate Swiss German more holistically, likely due to limited familiarity with internal dialect variation (see Section 1.1).

The central and most novel finding of this study concerns H3 (phonetic distance). The hypothesis predicted that greater phonetic distance between speech varieties would correlate with increasingly negative evaluations. The results clearly support this prediction.

Across all analyses, phonetic distance emerged as the most robust and consistent predictor of evaluative ratings. In both multiple regression and GLM(M) frame-

works, greater phonetic distance reliably decreased the likelihood of positive ratings while increasing the likelihood of negative ratings.

Importantly, once phonetic distance is included in the models, the former explanatory power of region is substantially reduced or in parts completely erased from statistical significance: the effect operates with similar strength across all regions investigated. Even among Swiss participants evaluating Swiss German variants, phonetic distance predicts ratings just as strongly as it does among German participants evaluating the same material.

This opens the question concerning phonological systematicity. While seemingly unexpected, our findings actually support earlier claims by Schmitt (1992), Kehrein (2015), and Kiesewalter (2019) that systemic, phonologically contrasting variants play a greater role in the perception and evaluation of regional speech. This is reflected in our weighted Levenshtein distance measurements, where systemic differences are assigned higher costs than non-systemic, purely phonetic deviations. Consequently, these contrasts contribute more strongly to higher phonetic distance values, which in turn elicit more negative ratings among all groups. These systemic variants likely also represent the most perceptually salient features, though we did not explicitly test salience in the present study. Next to effects of perceptual salience, it remains an open question whether the increasing phonetic distance observed in our data is associated with actual comprehension difficulties, as reported by Schmitt (1992), which might in turn drive negative evaluations. Future research should investigate the link between phonetic distance, salience, intelligibility, and social evaluation.

Taken together, these results highlight phonetic distance as a central factor shaping evaluative perception. Its effect not only operates across different listener groups but, in many cases, outweighs regional familiarity with the varieties being evaluated. This observation provides an important bridge to the final hypothesis, which addresses the role of familiarity and exposure in shaping evaluative behavior.

H4 predicted that the German rater groups, with their presumed lower familiarity with Swiss German varieties compared to Swiss participants, would exhibit stronger distance-related effects. Our results, however, show that the influence of phonetic distance is equally robust across all regions. Consequently, H4 must be rejected.

This finding carries broader implications: although we observe systematic differences in pre-existing attitudes toward regionally marked speech, with Swiss participants showing the expected higher dialect affinity than German participants, these predispositions do not diminish the strong, consistent effect of phonetic distance across all rater groups. In other words, even though H2 is confirmed (systematic attitudinal differences exist) the structural phonetic differences between varieties seem to override these predispositions in shaping evaluative behavior. Consequently, the earlier suggestion in this section that the differing sociolinguistic situations of the regions involved explain rating behavior appears less compelling.

This raises another central question in the study of speech perception concerning the relationship between phonetic distance and listener attitudes: Can these factors be treated as independent variables, or does phonetic variation, as one of several components, rather actively shape attitudes toward speech variation? This perspective best aligns with flexible, constructivist frameworks of language attitude models, such as those proposed by Purschke (2015) and Soukup (2019). As a theoretical extension of such models, language attitudes would have to be modeled as emergent from multiple interacting components, which may in turn draw on several attitude objects involving different cognitive processes, as in the present study.

Ultimately, the distinction between purely linguistic and social factors in explanatory models of linguistic evaluation and social judgment seems to need revision. The consensus of recent sociophonetic research is clear: social factors can shape listeners' phonetic judgments, while phonetic variation simultaneously influences social evaluations (cf. Kendall et al. 2023: 10). However, in light of our findings, critically re-evaluating the relative weight of social factors including speaker background, and purely linguistic or structural factors in speech perception and evaluation, appears desirable for the field of sociophonetics, at least within the DACH-region.

From a theoretical perspective, the present findings allow us to rethink the relationship between inherent value and social connotations (Giles et al. 1979; Giles &

Niedzielski 1998). While the results clearly show that phonetic distance, conceived of as a purely linguistic measure, outweighs attitudinal predispositions as socially connotated constructs in predicting evaluative behavior, it remains questionable to what extent such effects can be considered independent of an overall social meaning. As operationalized here, phonetic distance ultimately captures the degree to which a variety differs from a listener's own habitual speech pattern. In this sense, stronger negative evaluations of more distant variants may reflect a general preference for linguistic similarity, aligning with findings such as those of Heblich et al. (2015), where linguistic proximity enhances trust in economic decision-making. From this perspective, what appears as an inherent-value effect may, at least in part, be grounded in socially mediated processes and even structurally defined differences may derive their evaluative force from their role in social interaction.

Despite the limitations already noted above, several constraints of this study should be considered, particularly regarding the generalizability of the findings. First, although phonetic distance consistently predicted positive or negative judgments across all rater groups, this trend must be interpreted cautiously.

The relatively small and imbalanced sample size may limit statistical power and thus the robustness of the models. At the same time, the analyses revealed clear, consistent and statistically significant trends, particularly with respect to phonetic distance, suggesting that the observed patterns are unlikely to be merely artefactual.

It should however be noted that the models account for only a limited proportion of the variance (approximately 6–7% in the attitudes-plus-region models and around 12% in the phonetic-distance-plus-region models), indicating that a substantial share of the variation remains unexplained. This suggests that additional factors such as item-specific effects shown for example in Pistor and Leemann (2024), individual listener characteristics and personality traits (cf. Winkler et al. 2023), other unmeasured social variables (e.g., age or gender, cf. Bülow et al. 2021) and lastly, dialect identification, may also play a role in the evaluation process of linguistic input. Concerning the latter, depending on the listeners' ability to identify a given variety, associated social meanings might be activated (cf. the social connotations hypothesis).

Furthermore, the study was based on isolated words produced in a laboratory setting, without contextual information, and the rater sample was relatively homogeneous in age. As has been shown, combinations of linguistic variables can index divergent social meanings, and the indexical value of features may shift across generations (cf. Pharao 2025: 359).

Finally, attitudinal measures addressed dialects in general rather than the specific varieties under investigation. Consequently, it remains uncertain how explicit attitudes toward Bern and Zurich German might have influenced judgments or the modeling of attitudes as predictors, particularly among German participants from Saxony and Hesse. This limitation is partly mitigated by prior research on Swiss Germans' evaluations of these varieties, but comparable data for the German participants are largely lacking. However, the design was unidirectional: all stimuli were produced by Swiss speakers, whereas listeners came from both Switzerland and Germany. Whether the observed role of phonetic distance holds in the reverse direction, i.e., for Swiss listeners evaluating speakers from Hesse or Saxony, therefore remains open.

Future research should address these dimensions more systematically. At the same time, it must be emphasized that the aim of the present study was not to account for the full range of variation, but rather to explore the role of phonetic distance as a comparatively neglected factor alongside more commonly examined attitudinal predictors. While prior research emphasizes that the relevance of linguistic and social cues often depends on the listener (cf. Kendall et al. 2023: 10–12), the present effect of phonetic distance appeared robust across groups differing in familiarity, similarity, and social distance, and might open doors to further sociophonetic investigation.

Author Statement

The original idea took shape during joint discussions among the three authors. TP, AL, and BS collected the data. TP developed the study design and methodology, conducted the transcriptions and measurements, curated and analyzed the data, fitted all models,

created the visualizations, and wrote all drafts of the manuscript. AL and BS validated the transcriptions and commented on the drafts, which TP implemented. TP carried out the revisions, which were discussed with and agreed upon by AL and BS.

Data availability

Experiment data, R-code and stimuli are available on University of Bern's Open Research Repository: https://boris-portal.unibe.ch/handle/20.500.12422/254916

Endnotes

1 It has been suggested that segment-based approaches might more aptly be termed symbol-based, since segments can be analyzed acoustically as well as through auditory assessment. We acknowledge this fact, but retain the term segment-based, as opposed to signal-based approaches, for reasons of international comparability: it is the term used throughout the literature cited here on this topic (Lind-Combs et al. 2023; Wieling et al. 2014; Heeringa 2004).

References

Adler, Astrid & Albrecht Plewnia. 2018. Möglichkeiten und Grenzen der quantitativen Spracheinstellungsforschung. In Lenz, Alexandra N. & Albrecht Plewnia (eds.). *Variation – Normen – Identitäten*, 63–97. Berlin: De Gruyter.

Adler, Astrid & Albrecht Plewnia. 2019. Die Macht der großen Zahlen. Aktuelle Spracheinstellungen in Deutschland. In Eichinger, Ludwig M. & Albrecht Plewnia (eds.). *Neues vom heutigen Deutsch. Empirisch – methodisch – theoretisch*, 141–162. Berlin: De Gruyter.

Adler, Astrid & Albrecht Plewnia. 2020. Aktuelle Bewertungen regionaler Varietäten des Deutschen. In Hundt, Markus, Andrea Kleene, Albrecht Plewnia & Verena Sauer (eds.). *Regiolekte. Objektive Sprachdaten und subjektive Sprachwahrnehmung*, 15–36. Tübingen: Narr.

Ammon, Ulrich. 1995. *Die deutsche Sprache in Deutschland, Österreich und der Schweiz. Das Problem der nationalen Varietäten*. Berlin: De Gruyter.

Anders, Christina A. 2010a. *Wahrnehmungsdialektologie: das Obersächsische im Alltagsverständnis von Laien*. Berlin: De Gruyter.

Anders, Christina A. 2010b. Die wahrnehmungsdialektologische Rekodierung von laienlinguistischem Alltagswissen. In Anders, Christina A., Markus Hundt & Alexander Lasch (eds.). *Perceptual Dialectology. Neue Wege der Dialektologie*, 67–89. Berlin: De Gruyter.

Anikin, Andrey, Nikolay Aseyev & Niklas E. Johansson. 2023. Do some languages sound more beautiful than others? *Psychological and Cognitive Sciences* (*PNAS*) 120(17). e22183671202023.

Arendt, Birte. 2019. Wie sagt man hier? Bewertungen von Dialekt, Regionalsprache und Standard im Spannungsfeld regionaler Identität und sozialer Distinktion. In Antos, Gerd, Thomas Niehr & Jürgen Spitzmüller (eds.). *Handbuch Sprache im Urteil der Öffentlichkeit*, 333–352. Berlin: De Gruyter.

Aryani, Arash, Markus Conrad, David Schmidtke & Arthur Jacobs. 2018. Why 'piss' is ruder than 'pee'? The role of sound in affective meaning making. *PLoS ONE* 13(6). e0198430. https://doi.org/10.1371/journal.pone.0198430

Auer, Peter. 2005. Europe's sociolinguistic unity, or: A typology of European dialect/standard constellations. In Delbeque, Nicole, Johan van der Auwera & Dirk Geeraerts (eds.). *Perspectives on Variation*, 7–42. Berlin: De Gruyter.

Bartelds, Martijn, Caitlin Richter, Mark Liberman & Martijn Wieling. 2020. A New Acoustic-Based Pronunciation Distance Measure. *Frontiers in Artificial Intelligence* 3(39). https://doi.org/10.3389/frai.2020.00039

Berthele, Raphael. 2010. Der Laienblick auf sprachliche Varietäten. Metalinguistische Vorstellungswelten in den Köpfen der Deutschschweizerinnen und Deutschschweizer. In Anders, Christina A., Markus Hundt & Alexander Lasch (eds.). *Perceptual Dialectology. Neue Wege der Dialektologie*, 245–267. Berlin: De Gruyter.

Bülow, Lars, Andrin Büchler, Nicolai Rawyler, Christa Schneider & David Britain. 2021. Linguistic, Social, and Individual Factors Constraining Variation in Spoken Swiss Standard German. In Werth, Alexander, Lars Bülow, Simone E. Pfenninger & Markus Schiegg (eds.). *Intra-individual Variation in Language*, 127–174. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110743036-005

Christen, Helen. 2010. Was Dialektbezeichnungen und Dialektattribuierungen über alltagsweltliche Konzeptualisierungen sprachlicher Heterogenität verraten. In Anders, Christina A., Markus Hundt & Alexander Lasch (eds.). *Perceptual Dialectology. Neue Wege der Dialektologie*, 269–291. Berlin: De Gruyter.

Christen, Helen. 2019. Alemannisch in der Schweiz. In Herrgen, Joachim & Jürgen E. Schmidt (eds.). *Sprache und Raum. Ein internationales Handbuch der Sprachvariation. Band 4: Deutsch*, 246–279. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110261295-009

Christen, Helen & Regula Schmidlin. 2019. Die Schweiz. Dialektvielfalt in mehrsprachigem Umfeld. In Beyer, Rahel & Albrecht Plewnia (eds.). *Handbuch des Deutschen in West- und Mitteleuropa. Sprachminderheiten und Mehrsprachigkeitskonstellationen*, 193–244. Tübingen: Narr.

Clopper, Cynthia G. 2024. Dynamic acoustic vowel distances within and across dialects. *Journal of the Acoustical Society of America* 156. 2497–2507. https://doi.org/10.1121/10.0032385

Coupland, Nikolas & Hywel Bishop. 2007. Ideologised values for British accents. *Journal of Sociolinguistics* 11(1). 74–93. https://doi.org/10.1111/j.1467-9841.2007.00311.x

Eichinger, Ludwig M., Albrecht Plewnia, Christiane Schoel & Dagmar Stahlberg (eds.). 2012. *Sprache und Einstellungen. Spracheinstellungen aus sprachwissenschaftlicher und sozialpsychologischer Perspektive*. Tübingen: Narr.

Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12. 453–476. https://doi.org/10.1111/j.1467-9841.2008.00374.x

Flege, James E. 1984. The detection of French accent by American listeners. *Journal of the Acoustical Society of America* 76(3). 692–707. https://doi.org/10.1121/1.391256

Fuertes, Jairo N., William H. Gottdiener, Helena Martin, Tracey C. Gilbert & Howard Giles. 2012. A meta-analysis of the effects of speakers' accents on interpersonal evaluations. *European Journal of Social Psychology* 42(1). 120–133. https://doi.org/10.1002/ejsp.862

Ganswindt, Brigitte, Roland Kehrein, Carolin Kiesewalter & Tillmann Pistor. 2023. *Regionalakzente in Deutschland, Östliches Mitteldeutsch.* Online. https://dsa.info/regionalakzente/mdt_ost.html [last accessed 5/3/2026].

Giles, Howard, Richard Bourhis & Ann Davis. 1979. Prestige Speech Styles: The Imposed Norm and Inherent Value Hypotheses. In McCormack, William C. & Stephen A. Wurm (eds.). *Language and Society: Anthropological Issues*, 589–596. The Hague: Mouton.

Giles, Howard & Nancy Niedzielski. 1998. Italian is beautiful, German is ugly. In Bauer, Laurie & Peter Trudgill (eds.). *Language Myths*, 85–93. London: Penguin.

Greenwald, Anthony G., Brian A. Nosek & Mahzarin R. Banaji. 2003. Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology* 85(2). 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Heblich, Stephan, Alfred Lameli & Gerhard Riener. 2015. The Effect of Perceived Regional Accents on Individual Economic Behavior: A Lab Experiment on Linguistic Performance, Cognitive Ratings and Economic Decisions. *PLoS ONE* 10(2). e0113475. https://doi.org/10.1371/journal.pone.0113475

Heeringa, Wilbert. 2004. *Measuring Dialect Pronunciation Differences using Levenshtein Distance*. Groningen: University of Groningen. (doctoral thesis)

Hengartner, Thomas. 1993. Faktoren der Einschätzung Schweizerdeutscher Dialekte. In Lüdi, Georges & Claude Anna Zuber (eds.). *Contributions aux 4èmes rencontres régionales de linguistique, Bâle 1992*, 127–141. Basel: Romanisches Seminar der Universität Basel.

Herrgen, Joachim & Jürgen E. Schmidt. 1989. Dialektalitätsareale und Dialektalitätsabbau. In Putschke, Wolfgang, Werner H. Veith & Peter Wiesinger (eds.). *Dialektgeographie und Dialektologie: Günter Bellmann zum Geburtstag von seinen Schülern und Freunden*, 304–346. Marburg: Elwert.

Herrgen, Joachim, Alfred Lameli, Stefan Rabanus & Jürgen E. Schmidt. 2001. *Dialektalität als phonetische Distanz. Ein Verfahren zur Messung standarddivergenter Sprechformen*. Marburg: Online. http://archiv.ub.uni-marburg.de/es/2008/0007/pdf/dialektalitaetsmessung.pdf [last accessed 5/3/2026].

Hettler, Yvonne. 2018. *Salienz, Bewertung und Realisierung regionaler Sprachmerkmale in Bremen und Hamburg*. Hildesheim: Olms.

Hotzenköcherle, Rudolf & Heinrich Baumgartner (eds.). 1962–1997. *Sprachatlas der deutschen Schweiz*. Bern: Francke.

Hundt, Markus. 2010. Bericht über die Pilotstudie „Laienlinguistische Konzeptionen deutscher Dialekte". In Anders, Christina A., Markus Hundt & Alexander Lasch (eds.). *Perceptual Dialectology. Neue Wege der Dialektologie*, 179–215. Berlin: De Gruyter.

Hundt, Markus. 2012. Warum gibt es eigentlich „beliebte" und „unbeliebte" Dialekte? Theorien und Methoden der Einstellungsforschung im Bereich der Wahrnehmungsdialektologie. In Hünecke, Rainer & Karlheinz Jakob (eds.). *Die obersächsische Sprachlandschaft in Geschichte und Gegenwart*, 175–222. Heidelberg: Winter.

Hundt, Markus. 2017. Struktur und Komplexität des linguistischen Laienwissens. In Hundt, Markus, Nicole Palliwoda & Saskia Schröder (eds.). *Der deutsche Sprachraum aus der Sicht linguistischer Laien. Ergebnisse des Kieler DFG-Projektes*, 121−159. Berlin: De Gruyter.

Hundt, Markus, Nicole Palliwoda & Saskia Schröder. 2015. Wahrnehmungsdialektologie – Der deutsche Sprachraum aus der Sicht linguistischer Laien. Exemplarische Ergebnisse des Kieler DFG-Projekts. In Kehrein, Roland, Alfred Lameli & Stefan Rabanus (eds.). *Regionale Variation des Deutschen. Projekte und Perspektiven*, 585–620. Berlin: De Gruyter.

Kehrein, Roland. 2012. *Regionalsprachliche Spektren im Raum. Zur linguistischen Struktur der Vertikale*. Stuttgart: Steiner.

Kehrein, Roland. 2015. Deutsche Regionalakzente – ihre Entstehung, Form und mögliche Weiterentwicklung. In Elmentaler, Michael, Markus Hundt & Jürgen E. Schmidt (eds.). *Deutsche Dialekte. Konzepte, Probleme, Handlungsfelder. Akten des 4. Kongresses der Internationalen Gesellschaft für Dialektologie des Deutschen*, 453–477. Stuttgart: Steiner.

Kehrein, Roland. 2019. Areale Variation im Deutschen „vertikal". In Herrgen, Joachim & Jürgen E. Schmidt (eds*.). Sprache und Raum. Ein internationales Handbuch der Sprachvariation. Band 4: Deutsch*, 121–158. Berlin: De Gruyter Mouton.

Kendall, Tyler, Nicolai Pharao, Jane Stuart-Smith & Charlotte Vaughn. 2023. Advancements of phonetics in the 21st century: Theoretical issues in sociophonetics. *Journal of Phonetics* 98(2023). 101226. https://doi.org/10.1016/j.wocn.2023.101226

Kiesewalter, Carolin. 2019. *Zur subjektiven Dialektalität regiolektaler Aussprachemerkmale des Deutschen*. Stuttgart: Steiner.

Knops, Uus. 1988. Attitudes towards Regional Variation in Dutch Pronunciation. In van Hout, Roeland and Uus Knops (eds.). *Language Attitudes in the Dutch Language Area*, 105–120. Berlin: De Gruyter Mouton.

Kogan, Vita V. & Reiterer, Susanne M. 2021. Eros, Beauty, and Phon-Aesthetic Judgments of Language Sound. We Like It Flat and Fast, but Not Melodious. Comparing Phonetic and Acoustic Features of 16 European Languages. *Frontiers in Human Neuroscience* 15(578594).

Lambert, Wallace E., Richard C. Hodgson, Robert Gardner & Samuel Fillenbaum. 1960. Evaluational Reactions to Spoken Languages. *Journal of Abnormal and Social Psychology* 60. 44–51.

Lameli, Alfred. 2004. *Standard und Substandard. Regionalismen im diachronen Längsschnitt.* Stuttgart: Steiner.

Lameli, Alfred. 2013. *Strukturen im Sprachraum. Analysen zur arealtypologischen Komplexität der Dialekte in Deutschland*. Berlin: De Gruyter.

Lameli, Alfred, Elvira Glaser & Philipp Stoeckle. 2020. Drawing areal information from a corpus of noisy dialect data. *Journal of Linguistic Geography* 8(1). 31–48.

Lameli, Alfred. 2025. Gesprochenes Deutsch in den Regionen. Eine Standortbestimmung für die Bundesrepublik Deutschland. In Proske, Nadine, Thilo Weber, Monika Dannerer & Arnulf Deppermann (eds.). *Gesprochenes Deutsch. Struktur, Variation, Interaktion*, 51–80. Berlin: De Gruyter.

Leemann, Adrian, Marie-José Kolly & Francis Nolan. 2015. It's not phonetic aesthetics that drives dialect preference: The case of Swiss German. In *Proceedings of International Conference on Phonetic Sciences 2015*.

Levenshtein, Vladimir I. 1966. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics-Doklady* 10(8). 707–710.

Levy, Helena, Lars Konieczny & Adriana Hanulíková. 2019. Processing of unfamiliar accents in monolingual and bilingual children: effects of type and amount of accent experience. *Journal of Child Language* 46(2). 368–392. https://doi.org/10.1017/S030500091800051X

Lind-Combs, Holly C., Tessa Bent, Rachael F. Holt, Cynthia G. Clopper & Emma Brown. 2023. Comparing Levenshtein distance and dynamic time warping in predicting listeners' judgments of accent distance. *Speech Communication* 155 (2023). 102987. https://doi.org/10.1016/j.specom.2023.102987

Montgomery, Chris & Emma Moore. 2018. Evaluating S(c)illy voices: The effects of salience, stereotypes, and co-present language variables on real-time reactions to regional speech*. Language* 94(3). 629–661.

Mooshammer, Christine, Dominique Bobeck, Henrik Hornecker, Kierán Meinhardt, Olga Olina, Marie Christin Walch & Qiang Xia. 2023. Does Orkish Sound Evil? Perception of Fantasy Languages and Their Phonetic and Phonological Characteristics. *Language and Speech* 67(4). 961–1000.

Park, Hanyong. 2013. Detecting foreign accent in monosyllables: the role of L1 phonotactics. *Journal of Phonetics* 41. 78–87.

Pharao, Nicolai. 2025. Language Attitudes and Language Change. In Cieri, Christopher, Lauren Hall-Lew, Katie Drager & Malcah Yaeger-Dror (eds). *Dimensions of Linguistic Variation*, 350–372. Oxford: Oxford University Press.

Pharao, Nicolai & Tore Kristiansen. 2019. Reflections on the relation between direct/indirect methods and explicit/implicit attitudes. *Linguistics Vanguard* 5(s1). 20180010. https://doi.org/10.1515/lingvan-2018-0010

Pistor, Tillmann, Hanni Schnell & Lars Vorberger. 2023. *Regionalakzente in Deutschland, Westliches Mitteldeutsch*. Online. https://dsa.info/regionalakzente/mdt_west.html [last accessed 5/3/2026].

Pistor, Tillmann & Adrian Leemann. 2024. Echoes of Implicit Bias. Exploring Aesthetics and Social Meanings of Swiss German Dialect Features. In *Interspeech 2024*, 447–451.

Plewnia, Albrecht & Astrid Rothe. 2012. Sprache – Einstellungen – Regionalität. In Eichinger, Ludwig M., Albrecht Plewnia, Christiane Schoel & Dagmar Stahlberg (eds.). *Sprache und Einstellungen. Spracheinstellungen aus sprach-wissenschaftlicher und sozialpsychologischer Perspektive,* 9–118. Tübingen: Narr.

Preston, Dennis R. 2019. Folk Linguistics and the Perception of Language Variety. In Antos, Gerd, Thomas Niehr & Jürgen Spitzmüller (eds.). *Handbuch Sprache im Urteil der Öffentlichkeit*, 140–164. Berlin: De Gruyter.

Purschke, Christoph. 2011. *Regionalsprache und Hörerurteil. Grundzüge einer perzeptiven Variationslinguistik*. Stuttgart: Steiner.

Purschke, Christoph. 2015. REACT – A constructivist theoretic framework for attitudes. In Prikhodkine, Alexei & Dennis R. Preston (eds.). *Responses to Language Varieties: Variability, processes and outcomes*, 37–54. Amsterdam: John Benjamins.

Purschke, Christoph. 2020. Attitudes Toward Multilingualism in Luxembourg. A Comparative Analysis of Online News Comments and Crowdsourced Questionnaire Data. *Frontiers in Artificial Intelligence* 3. 536086. https://doi.org/10.3389/frai.2020.536086

Rakić, Tamara, Melanie C. Steffens & Amélie Mummendey. 2011. Blinded by the accent! The minor role of looks in ethnic categorization. *Journal of Personality and Social Psychology* 100(1). 16–29. https://doi.org/10.1037/a0021522

Reiterer, Susanne Maria, Vita V. Kogan, Annemarie Seither-Preisler & Gašper Pesek. 2020. Foreign language learning motivation: phonetic chill or latin lover effect? Does sound structure or social stereotyping drive FLL? *Psychology of Learning and Motivation* 72. 165–205.

Ris, Roland. 1992. Innerethnik der deutschen Schweiz. In Hugger, Paul (ed.). *Handbuch der schweizerischen Volkskultur*, Vol. II, 749–766. Zürich: Offizin.

Rocholl, Marie Josephine. 2015. *Ostmitteldeutsch – eine modern Regionalsprache? Eine Untersuchung zu Konstanz und Wandel im thüringisch-obersächsischen Sprachraum*. Hildesheim: Olms.

Rosenberg, Milton J. & Carl I. Hovland. 1960. Cognitive, affective and behavioral components of attitudes. In Rosenberg, Milton J. & Carl I. Hovland (eds). *Attitude organization and change: An analysis of consistency among attitude components*, 1–14. New Haven, CT: Greenwood.

Schaufuß, Anja & Beat Siebenhaar. 2012. Spracheinstellungen und phonetische Variation als Ausdruck verschwommener Dialektabgrenzung. *Zeitschrift für Literaturwissenschaft und Linguistik* 166. 88–109. https://doi.org/10.1007/BF03379865

Schaufuß, Anja. 2015. Regionalsprachlichkeit von Sprechern des Obersächsischen im Dreieck Dresden, Chemnitz, Leipzig. Hörerurteil und phonetischer Abstand im diatopischen Vergleich. In Elmentaler, Michael, Markus Hundt & Jürgen E. Schmidt (eds.). *Deutsche Dialekte. Konzepte, Probleme, Handlungsfelder. Akten des 4. Kongresses der Internationalen Gesellschaft für Dialektologie des Deutschen (IGDD)*, 359–377. Stuttgart: Steiner.

Scherrer, Yves & Stoeckle, Philipp. 2016. A quantitative approach to Swiss German – Dialectometric analyses and comparisons of linguistic levels. *DiG* 2016(24). 92–125. https://doi.org/10.1515/dialect-2016-0006

Schmidt, Jürgen E. 2017. Vom traditionellen Dialekt zu den modernen deutschen Regionalsprachen. In Deutsche Akademie für Sprache und Dichtung, Union der deutschen Akademien der Wissenschaften (ed.). *Vielfalt und Einheit der deutschen Sprache. Zweiter Bericht zur Lage der deutschen Sprache*, 103–141. Tübingen: Stauffenburg.

Schmitt, Ernst H. 1992. *Interdialektale Verstehbarkeit. Eine Untersuchung im Rhein- und Moselfränkischen*. Stuttgart: Steiner.

Schröder, Saskia. 2019. *Sprachräumliche Praxis. Sprachraumkartierung in der Wahrnehmungsdialektologie*. Berlin: Lang.

Sharma, Devyani, Erez Levon, & Yang Ye. 2022. 50 years of British accent bias: Stability and lifespan change in attitudes to accents. *English World-Wide* 43(2). 135–166. https://doi.org/10.1075/eww.20010.sha

Siebenhaar, Beat. 2012. Akkommodation und Sprachattitüden von Deutschschweizer Migranten im ostmitteldeutschen Sprachraum. *Sociolinguistica* 26(1). 151–164. https://doi.org/10.1515/soci.2012.26.1.151

Siebenhaar, Beat. 2014. Instrumentalphonetische Analysen zur Ausgestaltung des Sprechlagenspektrums in Leipzig. *Zeitschrift für Dialektologie und Linguistik* 2014(2). 151–190.

Soukup, Barabara. 2015. Mixing methods in the study of language attitudes: Theory and application. In Prikhodkine, Alexei & Dennis R. Preston (eds.). *Responses to language varieties: Variability, processes and outcomes*, 55–84. Amsterdam: John Benjamins.

Soukup, Barbara 2019. Sprachreflexion und Kognition: Theorien und Methoden der Spracheinstellungsforschung. In Antos, Gerd, Thomas Niehr & Jürgen Spitzmüller (eds*.). Handbuch Sprache im Urteil der Öffentlichkeit*, 83−106. Berlin: De Gruyter.

Steiner, Carina, Dieter Studer-Joho, Corinne Lanthemann, Andrin Büchler & Adrian Leemann. 2023. Sociodemographic and Attitudinal Effects on Dialect Speakers' Articulation of the Standard Language: Evidence from German-speaking Switzerland. In *Interspeech 2023*. 3542–3546.

Stoeckle, Philipp. 2014. *Subjektive Dialekträume im alemannischen Dreiländereck*. Hildesheim: Olms.

Stoeckle, Philipp & Christian Schwarz. 2019. Ethnodialektale Räume in der Deutschschweiz. In Nievergelt, Andreas & Ludwig Rübekeil (eds.). *Raum und Sprache. Festschrift für Elvira Glaser zum 65. Geburtstag*, 391–408. Heidelberg: Winter.

Van Bezooijen, René. 2002. Aesthetic evaluation of Dutch. In Long, Daniel & Dennis R. Preston (eds.). *Handbook of perceptual dialectology Vol. 2,* 13–30. Amsterdam: John Benjamins.

Vergeiner, Philip C. 2025. Dialect Classification and Everyday Culture: A Case Study from Austria. *Languages* 10(2). 17. https://doi.org/10.3390/languages10020017

Vō, Melissa; Markus Conrad, Lars Kuchinke, Karolina Urton, Markus Hofmann & Arthur Jacobs. 2009. The Berlin Affective Word List Reloaded (BAWL-R). *Behavior Research Methods* 41(2). 534–538.

Vorberger, Lars. 2019. *Regionalsprache in Hessen. Eine Untersuchung zu Sprachvariation und Sprachwandel im mittleren und südlichen Hessen*. Stuttgart: Steiner.

Vorberger, Lars. 2020. Rhein-Mainisierung – zur Neustrukturierung im hessischen Sprachraum. In Christen, Helen, Brigitte Ganswindt, Joachim Herrgen & Jürgen E. Schmidt (eds.). *Regiolekt – Der neue Dialekt? Akten des 6. Kongresses der Internationalen Gesellschaft für Dialektologie des Deutschen (IGDD),* 123–148. Stuttgart: Steiner.

Wesolek, Sarah, Piotr Gulgowski, Joanna Blaszczak & Marzena Zygis. 2024. The influence of L2 accent strength and different error types on personality trait ratings. In *Interspeech 2024.* 2–6.

Wieling, Martijn, Eliza Margaretha & John Nerbonne. 2014. Inducing a measure of phonetic similarity from pronunciation variation*. Journal of Phonetics* 40(2). 307–314.

Winkler, Anna, Vita V. Kogan & Susanne M. Reiterer. 2023. Phonaesthetics and personality – Why we do not only prefer Romance languages. *Frontiers in Language Sciences* 2. 1043619.