Quantifying dialectal input: Manual coding vs. perceptual ratings

Sarah Warchhold[1], *Bettina Braun*[1], *Katharina Zahner-Ritter*[2]

[1]*University of Konstanz*
[sarah.warchhold]/[bettina.braun]@uni-konstanz.de

[2]*University of Trier*
k.zahner-ritter@uni-trier.de

When investigating phonological representations in infants, accurately quantifying the phonological variability in children's input (e.g., induced by regional accents) can be challenging and time-consuming. Here, we test whether subjective ratings of perceived dialect strength, an efficient way of capturing dialectal variability, are a reliable and valid alternative to manually coding phonological alternations and hence quantify the variability in the input.
We sampled twenty 1-minute parental picture-descriptions of different dialect strengths (Braun et al. 2021). Perceived dialect strength was coded by 4 raters from different regions of Germany on a 4-point scale (cf. Floccia et al. 2009). Two independent annotators coded word forms in three ways: no variants, general variants or dialectal, using xml-snippets in Praat (for examples of coding of variants see Figure 1).

Reliability was high for the 4-point perceptual dialect scores (ICC=0.88) and moderate for the 3-point coding of word forms (dialectal, general, none, ICC=0.55). Dialect scores were averaged and disagreements in annotation of word forms were resolved. Resolved annotations were used for further analyses. To check whether perceptual ratings are a valid way of quantifying dialectal input, we correlated the mean proportion of general and dialectal variants in a sound file with mean dialect strength rating. Spearman's rho showed no correlation for general variants ($r=0.18$, $p>0.4$), but a strong positive correlation for the proportion of dialectal variants ($r=0.97$, $p<0.001$), see Figure 2.

In sum, our results showed high inter-rater reliability for perceptual judgments and a strong positive correlation with proportion of dialectal variants. Our data support the reliability of perceptual coding of dialect strength (cf. Grondelaers, van Hout & van der Harst 2015; Ryan 1973; Van Bezooijen & Van Hout 1985) revealing high correlation with the proportion of variant word forms, and hence seem a valid and reliable measure for quantifying dialectal input.

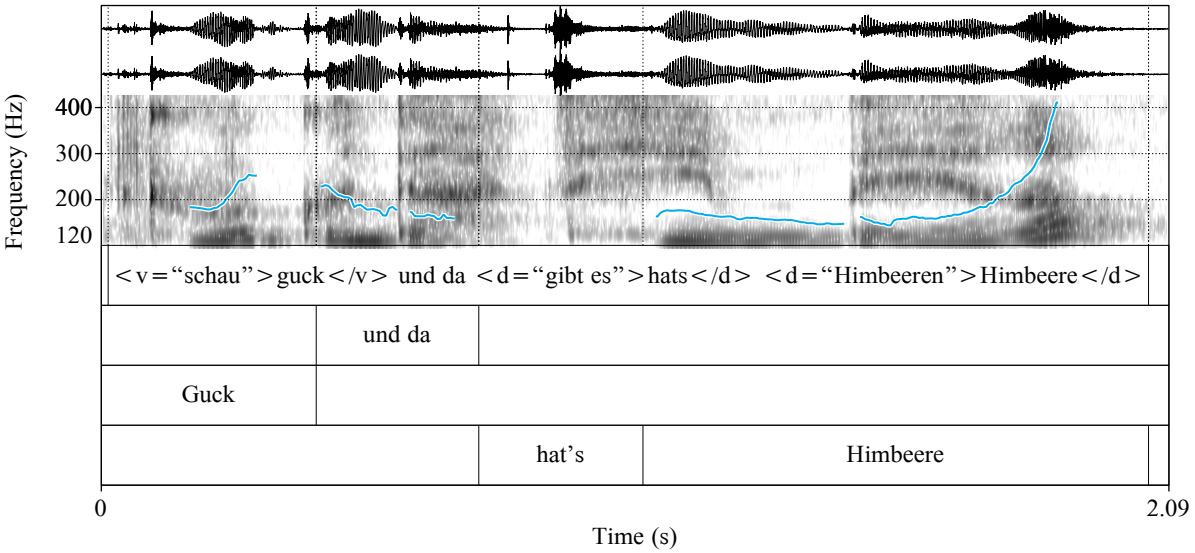

Figure 1 Exemplary excerpt from annotation process in Praat. Tier 1 shows the xml-snippets. Tier 2 shows word forms that were produced based on Standard German phonological rules. Tier 3 shows general variants and dialectal variants are shown in Tier 4. Clitics ('s) have been added to the preceding word and are not counted as an extra word. The file has a dialect strength score of 4. For this utterance, the proportion of general variants is 0.2 (1/5) and the proportion of dialectal variants is 0.4 (2/5).

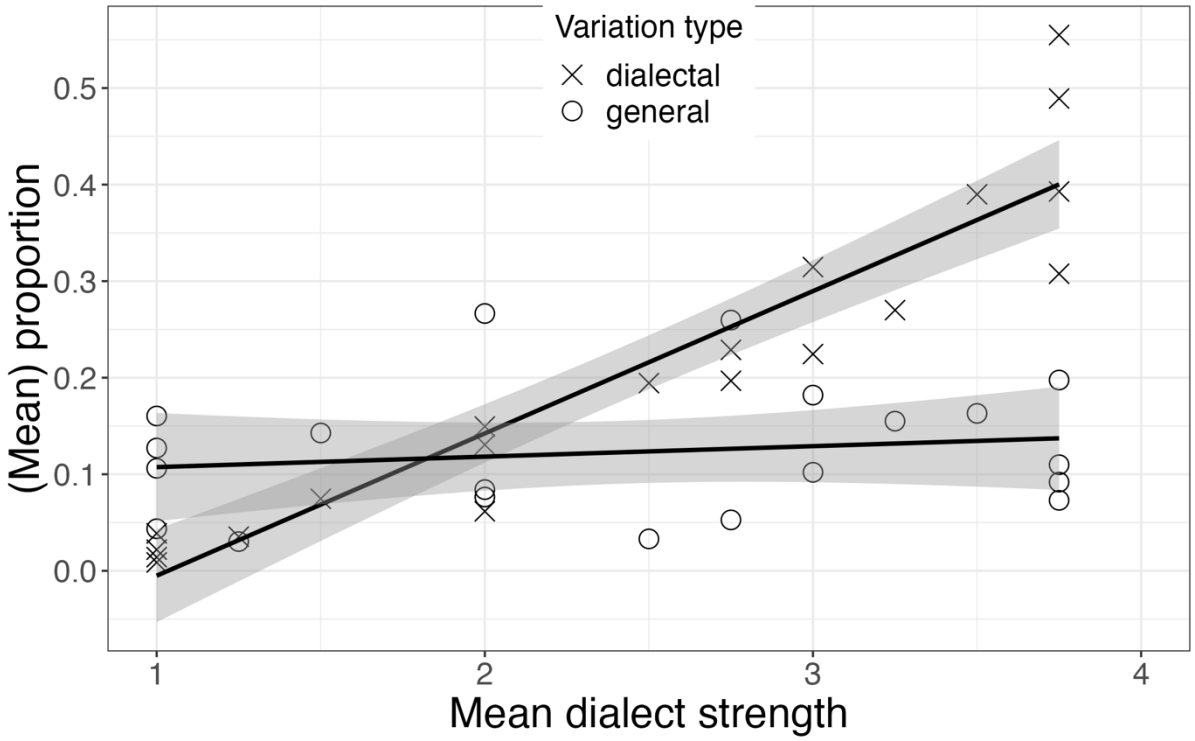

Figure 2 Scatterplot showing (mean) proportion of dialectal (X) and general (O) variants across mean dialect strength ratings and regression lines for dialectal variants starting at 0 and general variants starting at 0.1)

References

Braun, Bettina, Nathalie Czeke, Jasmin Rimpler, Claus Zinn, Robert Probst, Bastian Goldlücke, Julia Kretschmer & Katharina Zahner-Ritter (2021): Remote testing of the familiar word effect with non-dialectal and dialectal German-learning 1-2-year-olds. *Frontiers in Psychology* 714363.

Floccia, Caroline, J. Butler, Fréderique Girard & J. Goslin (2009): Categorization of regional and foreign accent in 5- to 7-year-old British children. *International Journal of Behavioral Development* 33(4) doi:10.1177/0165025409103871.

Grondelaers, Stefan, Roeland van Hout & Sander van der Harst (2015): Subjective accent strength perceptions are not only a function of objective accent strength. Evidence from Netherlandic Standard Dutch. *Speech Communication* 74. 1–11. doi:10.1016/j.specom.2015.07.004.

Ryan, E. B. (1973): Subjective reactions toward accented speech. In Roger W. Shuy & R. W. Fasold (Hrsg.), *Language Attitudes: Current Trens and Prospects*, 60–73. Washington, DC: Georgetown University Press.

Van Bezooijen, Renee & Roeland Van Hout (1985): Accentedness Ratings and Phonological Variables as Measures of Variation in Pronunciation. *Language and Speech* 28(2). 129–142. doi:10.1177/002383098502800203.