What influences the foreign accent strength? Phonological and grammatical errors in the perception of accentedness

Sarah Wesołek,[1,2], Piotr Gulgowski[1,3], Joanna Błaszczak[3], Marzena Żygis[1,2]

[1]Leibniz-Centre General Linguistics, Germany
[2]Humboldt University Berlin, Germany
[3] University of Wrocław, Poland

wesolek@leibniz-zas.de, piotr.gulgowski@uwr.edu.pl, zygis@leibniz-zas.de

Abstract

The present study investigates the influence of grammatical and phonological errors on the perceived degree of foreign accent strength. German and Polish participants listened to speech in their native language produced with foreign and native accent. They rated the accent strength of each sentence on a 7-point scale. Grammatical errors consisted of gender agreement violations and phonological errors consisted of controlled vowel substitutions.

Both error types significantly affected the perception of accent strength in the foreign and native-accented condition. In Polish, phonological anomalies had significantly more impact than grammatical violations in native-accented sentences. In German, there was no significant difference between phonological and grammatical violations.

The study provides evidence that the presence of phonological and grammatical errors increases the perceived accentedness of speech. The weighting of both errors for accent perception can vary between languages.

Index Terms: speech perception, foreign accent, grammatical violation, phonological substitution

1. Introduction

Late or incomplete language acquisition can lead to markedly foreign-sounding speech. Research shows that this can have significant consequences for spoken language processing and communication efficiency. Listening to non-native accented speech has been experimentally linked to a general processing slow-down [1, 2] and higher processing cost due to increased effort in speech perception [3]. As a result, listeners find it difficult to properly understand foreign-accented speakers, as shown experimentally in the form of lexical recognition problems [4, 5].

Accentedness is a dimension extending beyond narrowly defined language processing, with wider social and psychological consequences. If a person's speech is considered foreign, this can impact the way in which the speaker is perceived due to the presence of biases and stereotypes on the listener's side. Foreign pronunciation negatively influences the evaluation of the speaker's personality traits (see [6] for meta-analysis), and the truthfulness of their utterances [7]. As shown in [8], natives listening to foreign-accented speech tend to misjudge the speaker's affective state. Additionally, grammatical mistakes might be seen as a sign of generally low language skills. The negative associations linked to foreign accent can be strong enough to give rise to a form of illusion. There is emerging evidence [9] for the phenomenon of a 'grammatical tinnitus', where native listeners judge correct foreign-accented utterances more often to be ungrammatical than when produced by native speakers.

A vast body of research on foreign-accented speech indicates that the perceived accent strength correlates with listener-depended factors like the age of L2 language learning, the years of formal instruction, the length of residence in the second language environment, as well as the amount of L2 and L1 language use [10, 11]. Besides, it is suggested that accentedness can be influenced by several factors in the speech signal, like segmental substitutions, ungrammatical constructions, and deviations in the prosodic structure including intonation, timing, speech rate, and loudness, see [12, 13, 14, 15, 16, 17].

Less is known about the weighting of these specific features for the evaluation of a speaker's foreign accent strength by native listeners. Studies have shown that besides suprasegmental deviations, categorical substitutions can contribute to the notion of foreignness. When acquiring a new phonological system, foreign language learners tend to perceive similar sounds from the L2 as allophones of their native language [18, 19]. Consequently, they might substitute target language phonemes with similar sound categories from their native tongue. Categorical changes in phonemes have been shown to impact the recognition of non-native speech variants [13, 14], with the current state of research remaining somewhat unclear. For example, in a study on spontaneous and read speech produced by Japanese L2-English learners, the percentage of flap substitutions for the liquids /l/ and /r/ significantly interacted with the global foreign accent strength rated by native listeners [20]. The correlation, as interpreted by the authors, suggests that segmental violations are a strong contributor to foreign accent perception, although it is noteworthy that the data was not controlled for other possible accent indicators such as suprasegmental speech characteristics. Moreover, a recent study [21] assessed the relative influence of segmental (and suprasegmental) deviations on accentedness ratings in word-length stimuli spoken by four foreign speakers of different origins, compared to a native speaker of English. The study investigated the influence of relative Voice Onset Time (VOT) and vowel quality differences on global accentedness ratings. Whereas deviations in VOT predicted the perceived accent strength rated by native listeners towards words recorded by native Hindi, Spanish, Korean, and Mandarin speakers, changes in vowel quality correlated with the accentedness ratings of Korean and Mandarin speakers only. Similarly, [22] found longer VOTs in [p], [t], and [k], produced by Japanese learners to influence the perceived accent strength of sentences produced by foreign speakers.

A question not thoroughly investigated yet is whether language distortions that go beyond segmental and suprasegmental features have an impact on accent strength perception. The few available studies investigating the impact of grammatical errors on foreign accent ratings provide inconsistent results. [13] found the number of spontaneously produced grammatical errors to negatively influence the accent scores of utterances produced by foreign and native speakers. In contrast, when [23] manipulated the grammaticality of German sentences produced by native and foreign-accented speakers, they found that native listeners assigned higher accentedness ratings to native-accented speech containing grammatical errors, but not when sentences were produced by foreign speakers. The study suggests that grammatical violations have less impact on the perceived accent strength for sentences produced in foreign, compared to native-accented speech.

The present study aims to contrast the influence of phonological substitutions and grammatical errors on the perception of foreign accent strength in speech produced by native and foreign speakers. To our knowledge, none of the existing studies compared the influence of phonological versus grammatical errors in the perception of foreign accent strength directly. Based on the studies discussed above, as well as on our previous work [9], we hypothesize the following:

(H1) Phonological errors elicit higher accent strength ratings in comparison with the control condition and the grammatically incorrect condition across both, native and foreign-accented speech.

(H2) Grammatical errors elicit higher accent strength ratings in comparison with the control condition in native-accented, but not in foreign-accented speech, see [23].

We had no specific predictions regarding the potential differences between the German and Polish experiment.

2. Experiments

To test our hypotheses, we conducted two online experiments. One of them was conducted with German participants using German stimuli sentences. The other experiment was recorded with Polish participants using Polish stimuli. The experiments were matched in their experimental design. Participants listened to sentences produced by native and foreign speakers. Sentences produced by foreign speakers were marked by a subtle accent characterized by suprasegmental variation, apart from the deliberate phonological and grammatical manipulations described below.

Our experimental stimuli consisted of utterances that contained no explicit categorical or grammatical manipulations (control condition), sentences with a phonological substitution, or sentences with a grammatical error, leading to a 2 (native, foreign) x 3 (control, phonological substitution, grammatical error) design. The participant's task was to rate the foreign accent strength of the auditorily presented sentences on a seven-point scale from 1 (no accent) to 7 (strong accent). As a separate experimental task, participants also rated the speaker's personality traits. In the following, we limit ourselves to the discussion of the accent strength ratings.

2.1. Materials

For each experiment, 36 unique experimental items were created. Sentences of both experimental languages were of comparable length and followed the same syntactic structure.

We used general statements about various topics, mainly simple expressions about a person or object. For each item, three sentence versions (triples) were created: control sentences, sentences containing a phonological substitution, and sentences containing a grammatical error, resulting in 108 sentences. The three versions of each item differed only in the critical region which started with a preposition followed by a possessive determiner and a two- or three-syllabic noun. See Example 1 for a triple from the German experiment (*'Lena consults the teacher about her mistake in the exam'*) and Example 2 for a triple from the Polish experiment (*'Nina cleans up her flat before guest's arrival'*). For audio examples please visit the project's OSF repository (https://osf.io/k2mta/?view_only=f65bdededa9c4ad0b81c43c380ae5b3b). The phonologically anomalous condition contained a categorical vowel substitution that appeared in the stressed, penultimate syllable of the critical noun. In this condition, three native vowels were replaced in an equal number of sentences. We decided to use categorical vowel substitutions typical for the respective L2 learner group. Polish learners of German typically struggle with the German lax-tense contrast by substituting long tense German vowels with lax vowels from their native vowel inventory, see [24]. Germans, however, may tend to substitute the Polish lax vowels with tense vowels from their vowel inventory, when produced in stressed syllables. Additionally, we have chosen substitutions between [ʏ] and [ɨ], since [ʏ] does not belong to the Polish phoneme system and [ɨ] ist not part of the German vowel inventory. For German stimuli, the substitutions were: [e] → [ɛ], [o] → [ɔ], and [ʏ] → [ɨ]. Polish sentences contained the substitutions [ɛ]→[e], [ɔ]→[o], and [ɨ]→[ʏ]. The grammatical errors consisted of a gender mismatch between the possessive determiner and noun in the critical sentence region. Learners of languages with lexical gender often have problems with correct determiner-noun congruence, see [25].

Example 1: *Sentence triple (German experiment)*

Lena   befragt   die   Lehrerin ...
Lena   consults   the   teacher

(i)   control
    ...zu   ihrem   Fehler   in   der   Klausur.
    about   her[masc]   mistake[masc]   in   the   exam

(ii)   phonological substitution
    ... zu   ihrem   F[ɛ]ler   in   der   Klausur.
    about   her[masc]   mistake[masc]   in   the   exam

(iii)   grammatical error
    ... zu   ihrer   Fehler   in   der   Klausur.
    about   her[fem]   mistake[masc]   in   the   exam

Example 2: *Sentence triple (Polish experiment)*

Nina   robi   porządek...
Nina   makes   order

(i)   control
    ...w   swojej   kwaterze   przed   przyjazdem   gości.
    in   her[fem]   flat[fem]   before   arrival   guests

(ii)   phonological substitution
    ...w   swojej   kwat[eː]rze   przed   przyjazdem   gości.
    in   her[fem]   flat[fem]   before   arrival   guests

(iii)   grammatical error
    ...w   swoim   kwaterze   przed   przyjazdem   gości.
    in   her[masc]   flat[fem]   before   arrival   guests

Materials were digitally recorded with four bilingual male speakers that were selected on the basis of a pilot study. Two L1 speakers of Polish were also L2 speakers of German, and two native speakers of German were L2 speakers of Polish, resulting in the same four male speakers producing sentences for both the Polish and German experiment. Foreign-accented sentences were characterized by a subtle accent (without categorical changes apart from the phonologically anomalous condition). Each recording was inspected by at least two phonetically trained linguists for categorical changes and accent strength. If a file contained any other categorical violations than the intended one, it was re-recorded. The different versions of experimental items were distributed across three lists such that each participant listened to one version of a given item only. In addition to judging the accentedness, the task involved rating several personality traits of the speaker for each sentence (results not reported in the present paper). In order to avoid too many rating judgments for a single sentence occurrence, each sentence was presented twice. The accentedness rating was made on each sentence occurrence. Thus 36 sentences appeared on each list, which, due to the repetition, resulted in 72 trials.

2.2. Procedure

The experiment was part of the larger 'Grammatical tinnitus and its role in the perception of foreign language accent. A comparison of German and Polish (GRANITUS)' project. Participants took part in an EEG experiment first, which is not further discussed in this paper. One week after the EEG experiment, they were sent a link for the perceptual experiment. The experiment was prepared with the PsychoPy software (version 2021.2.3, [26]) and conducted online on the Pavlovia platform (https://pavlovia.org/). The experimental session started with written instructions appearing on the screen. Participants were asked to listen to automatically played sentences and to rate the speaker's accent strength and personality traits on five 7-point scales (four for the personality traits, one for accent strength) displayed on the screen. The accent strength scale encompassed points from 1 (not accented) to 7 (strongly accented). The scale presentation did not include numbers. The point on the very left indicated 'no accent', whereas the point on the very right indicated a 'strong accent'. Once the participant selected a point on all scales, a green button appeared at the bottom of the screen. By clicking on it, the participant proceeded to the next trial. The present paper is limited to the presentation of the accent strength ratings.

2.3. Participants

The participant group for the German experiment included 33 native speakers of German (17 women, 16 men) aged between 21 and 35 years (M=26, SD=3.8). They were students of the Philipps-Universität Marburg, except for eight participants who were non-students. None of the participants reported having a linguistic background or knowledge of Polish.

In the Polish experiment, 30 native speakers of Polish (16 women, 14 men) took part. They were mostly students of the University of Wrocław, aged between 21 and 31 years (M=23, SD=2.4). None of the students had a linguistic background nor reported knowledge of German.

2.4. Statistics

All statistical analyses were conducted with R (version 4.2.0 [27]). For each experiment (German and Polish), we fitted

identical linear mixed-effects models for ordinal response variables (package 'Ordinal', [28]) with the Accent Strength Rating [1(no accent) - 7(strong accent)] as the dependent variable. The fixed factors were Error Type [control, phonological substitution, grammatical error], Accent Type [native, foreign], and their interaction. The statistical model included Participant as random intercept with Error Type, Accent Type, as well as their interaction as its random slopes. We also included Sentence Triple (i.e., the three sentence versions of each experimental item) as a random intercept, with Error Type as its slope. Finally, the Speaker (the person recorded) was included as random intercept, with Error Type as its slope. Because of issues with correlation, Error Type was removed as a slope for the random structure of Speaker in both models. Additionally, because of issues with high correlation, the interaction between Error Type and Accent Type was removed from the random structure of Participant in the model for the Polish data. For multiple comparisons we used the emmeans function from the 'emmeans' package [29].

For the statistical analysis of the German experiment, 2376 datapoints were submitted. The dataset of the Polish experiment contained 2160 datapoints. No datapoints were excluded.

3. Results

3.1. German experiment

The analysis of the German data revealed a significant interaction between Accent and Error Type (z=-2.79, p<0.05).

We predicted that phonological errors will elicit higher accent strength ratings in comparison with the control condition and the grammatically incorrect condition across both, native and foreign-accented sentences (H1). Our results indicate that phonological errors elicited significantly higher accent strength ratings, compared to the control conditions in native (z=-7.24, p<.0001) and foreign-accented (z=-3.22, p<.005) sentences, see Figure 1. When we compared sentences with a phonological substitution to sentences with grammatical errors, the analysis revealed no significant difference in accent strength ratings for native-accented sentences nor for sentences produced by foreign speakers.
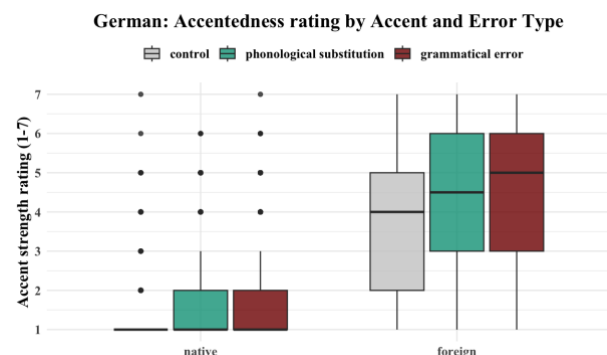

Figure 1: *Perceived foreign accent strength by Accent Type and Error Type (German Experiment)*

Furthermore, we hypothesized that grammatical errors will elicit higher accent strength ratings in comparison with the control condition in native-accented, but not in foreign-accented sentences (H2). Our results reveal that sentences containing grammatical errors were rated significantly higher than sentences in the control condition for both the native

(z=-6.10, p<.0001) and, contrary to our assumption, also in the foreign-accented condition (z=-3.93, p<.001).

3.2. Polish experiment

The statistical analysis of the Polish data revealed a significant interaction between Accent and Error Type (z=-2.35, p<.05).

In line with the first hypothesis, the Polish data revealed that phonological errors elicit significantly higher accent strength ratings compared to the control conditions in native (z=-6.78, p<.0001), and foreign-accented speech (z=-3.85, p<.001), see Figure 2. The analysis of Polish data, in contrast to the German results, confirmed that phonological substitutions had a significantly stronger impact than grammatical errors when produced in foreign (z=-2.46, p<.05), and in native-accented speech (z=-4.5, p<.0001).

Polish: Accentedness rating by Accent and Error Type

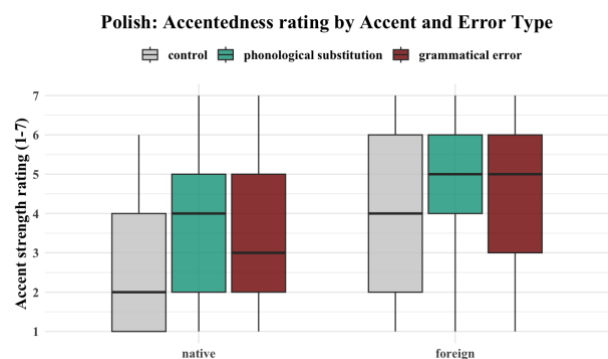control   phonological substitution   grammatical error

Figure 2: *Perceived foreign accent strength by Accent Type and Error Type (Polish Experiment)*

Regarding our second hypothesis, we predicted grammatical errors to influence the perceived accent strength in native-accented sentences but not in sentences uttered with a foreign accent. Like the German results, the Polish data showed that grammatical mistakes elicit higher accent strength ratings in comparison with the control condition in both foreign (z=2.79, p<.05) and native-accented sentences (z=-4.47, p<.0001).

4. Discussion and Conclusion

The present study explored the impact of phonological substitutions and grammatical violations on the perceived accent strength as rated by native listeners of Polish and German. Our design included sentences with (phonological) categorical vowel manipulations and (grammatical) determiner-noun incongruencies in contrast to sentences with no specific manipulations (control).

We predicted that sentences with phonological substitutions will elicit higher accent strength ratings when compared to sentences with no grammatical or phonological anomalies. Our first hypothesis also stated that phonological substitutions will have more impact on the perceived accent strength when compared to sentences with grammatical errors. The collected data indicates that within both languages phonological substitutions significantly influence the perceived foreign accent strength when embedded in native and foreign-accented speech. However, our hypothesis that phonological substitution will have a stronger impact on the perceived accentedness than grammatical errors, was confirmed only for Polish, but not for German. In Polish, regardless of the speaker's background (native, foreign), sentences received higher accentedness

ratings when they contained a vowel mismatch, compared to grammatical violations. In German, for both accents, no significant differences were found between the two error types.

Furthermore, we hypothesized grammatical errors to influence the perceived accent strength in sentences produced with native accent, but not in sentences produced by foreign-accented speakers. In contrast to [23], our analysis across the two languages indicated that grammatical mistakes influence the perceived accent strength in both foreign and native-accented sentences. The study, therefore, provides new empirical evidence that the presence of phonological anomalies and ungrammatical constructions can increase the perceived foreign-accentedness, when encountered in foreign and native speech. A question to explore in future research is to what extent the typicality of errors determines the strength of this effect. In our study, we used L2 vowel substitutions and grammatical gender disagreements, fairly typical errors commited by L2 learners. Less likely errors might impact the grammaticality ratings differently.

Another outcome of the present study is the difference in the relative importance of error types on the assessment of accent strength between German and Polish participants. Polish listeners were affected more by phonological anomalies than by grammatical violations, especially in the native accent condition. This is consistent with the fact that accentedness is a primarily phonological/phonetic phenomenon characterized by segmental and suprasegmental pronunciation problems. In contrast to Polish listeners, German participants assigned the same significance to grammatical errors and phonological anomalies when judging the speaker's accentedness. A possible explanation for this cross-linguistic discrepancy might lie in the much greater dialectal diversity and a possibly higher exposure to different foreign languages in German, compared to Polish raters. Due to a greater exposure to segmental variation between speakers of different dialects and foreign languages, German listeners might be less likely to consider segmental differences as sign of foreignness. Alternatively, the observed Polish-German divergence might stem from the differences in the direction of vowel substitutions picked for the two language groups. The lax-to-tense substitutions used for Polish stimuli might be more salient than the opposite substitutions used for German materials.

It is worth pointing out that the speakers providing recordings for the native Polish condition were considered by the native participants to be overall more foreign (even for the control sentences with no grammatical or phonological errors) than the speakers contributing their voices in the corresponding native German condition. It is hard to tell whether this is indicative of a systematic difference between foreign accent assessment between native speakers of Polish and German or due to the idiosyncratic features of our speaker and/or listener samples. The issue could be addressed in future studies.

5. Acknowledgements

6. References

[1] Clarke, C. M., Garrett, M. F. 2004. Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America*, 116 (6), 3647–3658.

[2] Floccia, C., Butler, J., Girard, F., Goslin, J. 2009. Categorization of regional and foreign accent in 5- to 7-year-old British children. *International Journal of Behavioral Development*, 33 (4), 366–375.

[3] Van Engen, K. J., Peelle, J. E. 2014. Listening effort and accented speech. *Frontiers in Human Neuroscience*, 8, 1–4.

[4] Bradlow, A. R., Pisoni, D. B. 1999. Recognition of spoken words by native and non-native listeners: talker-, listener-, and item-related factors, *The Journal of the Acoustical Society of America*, 106 (4), 2074–2085.

[5] Bent, T., Holt, R. F. 2013. The influence of talker and foreign-accent variability on spoken word identification, *The Journal of the Acoustical Society of America*, 133 (3), 1677–1686.

[6] Fuertes, J., Gottdiener, W., Martin, H., Gilbert, T., Giles, H. 2012. A meta-analysis of the effects of speakers' accents on interpersonal evaluations', *European Journal of Social Psychology*, 42, 120–133.

[7] Lev-Ari, S., Keysar, B. 2010. Less-Detailed Representation of Non-Native Language: Why Non-Native Speakers' Stories Seem More Vague, *Discourse Processes*, 49 (7), 523–538.

[8] Wesołek, S., Gulgowski, P., Błaszczak, J., Żygis, M. Illusions of Ungrammaticality in the Perception of Non-Native Accented Speech, *Proceedings of the International Congress of Phonetic Science*. Prague, August 7–11, 2023 (to appear).

[9] Holden, K.T., Hogan, J.T. 1993. The Emotive Impact of Foreign Intonation: An Experiment in Switching English and Russian Intonation. *Language and Speech*, 36, 67–88.

[10] Flege, J., Munro, M., Mackay, I. 1995. Effects of age of second-language learning on the production of English consonants', *Speech Communication*, 16, 1–26.

[11] Piske, T., MacKay, I. R. A., Flege, J. E. 2001. Factors affecting degree of foreign accent in an L2: a review, *Journal of Phonetics*, 29 (2), 191–215.

[12] Anderson-Hsieh, J., Johnson, R., Koehler, K. 1992. The Relationship Between Native Speaker Judgments of Nonnative Pronunciation and Deviance in Segmentals, Prosody, and Syllable Structure, *Language Learning*, 42 (4), 529–555.

[13] Munro, M. J., Derwing, T. M. 1995. Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. *Language Learning*, 45 (1), 73–97.

[14] Magen, H. S. 1998. The perception of foreign-accented speech. *Journal of Phonetics*, 26 (4), 381–400.

[15] de Mareüil, P. B., Vieru-Dimulescu, B. 2006. The Contribution of Prosody to the Perception of Foreign Accent. *Phonetica*, 63 (4), 247–267.

[16] Kang, O., 2010. Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness, *System*, 38 (2), 301–315.

[17] Oganyan, M., Wright, R., McCullough, E. 2021. Comparing Segmental and Prosodic Contributions to Speech Accent, in *Second Language Speech Learning: Theoretical and Empirical Progress*, Wayland, R., Ed. Cambridge: Cambridge University Press, 337–349.

[18] Flege, J. E. 1995. Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, 92, 233–277.

[19] Best, C. T., Tyler, M. D. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. in O.-S. Bohn & M. J. Munro, *Language Learning & Language Teaching,* 17, 13–34.

[20] Riney, T. J., Takada, M., Ota, M. 2000. Segmentals and Global Foreign Accent: The Japanese Flap in EFL', *TESOL Quarterly*, 34 (4), 711–737.

[21] Oganyan, M., Wright, R., McCullough, E. 2021. Comparing Segmental and Prosodic Contributions to Speech Accent, in *Second Language Speech Learning: Theoretical and Empirical Progress*, R. Wayland, Ed. Cambridge: Cambridge University Press, 337–349.

[22] Riney, T. J., & Takagi, N. 1999. Global foreign accent and voice onset time among Japanese EFL speakers. *Language Learning*, 49, 275–302.

[23] Asano, Y., Weber, A. 2016. Listener sensitivity to foreign-accented speech with grammatical errors. *Proceedings of the 38th Annual Conference of the Cognitive Science Society.* Austin, 1775–1780.

[24] Nimz, K., Khattab, G. 2020. On the role of orthography in L2 vowel production: The case of Polish learners of German. *Second Language Research*, 36 (4), 623–652.

[25] Grüter, T., Lew-Williams, C., Fernald, A. 2012. Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, 28 (2), 191–215.

[26] Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. 2019. PsychoPy2: experiments in behavior made easy. *Behavior Research Methods.*

[27] R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Version 4.2.0. https://www.R-project.org/.

[28] Christensen, R. H. B. 2022. ordinal—Regression Models for Ordinal Data. R package version 2022. 11–16.

[29] Lenth, R. 2022. emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.8.3. https://CRAN.R-project.org/package=emmeans.