

# Illusions of Ungrammaticality in the Perception of Non-Native Accented Speech

Sarah Wesolek<sup>1,2</sup>, Piotr Gulgowski<sup>3</sup>, Joanna Błaszczak<sup>3</sup>, Marzena Żygis<sup>1,2</sup>

<sup>1</sup>Leibniz-Centre General Linguistics, <sup>2</sup>Humboldt University Berlin, <sup>3</sup>University of Wrocław

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

## ABSTRACT

Studies have shown a reduced sensitivity to grammatical violations in non-native accented speech. The same expectation of ungrammaticality that reduces sensitivity to errors might also prime listeners to perceive illusory errors in accented utterances. In our study, we hypothesized that listeners exposed to non-native accented speech perceive more illusory grammatical errors. This effect should increase when sentences contain phonological violations. To test our predictions, we conducted two mirror experiments for German and Polish, contrasting the perception of: (a) well-formed, (b) phonologically anomalous, and (c) ungrammatical sentences in native and non-native speech.

Our results show that in both languages grammatical sentences were more often judged as ungrammatical in non-native than native accent, a phenomenon called ‘grammatical tinnitus’. However, the effect did not increase when phonologically anomalous sentences were presented in the Polish experiment, even if the overall number of incorrect answers increased. For German, the ‘tinnitus’ was absent in this condition.

**Keywords:** non-native accent, perception, grammatical illusion, German, Polish

## 1. INTRODUCTION

The current study assesses whether native listeners perceive illusions of ungrammaticality when listening to foreign-accented speech. The investigated phenomenon can be illustrated with the real-life example of a German politician, Raed Saleh, who was born in Palestine and moved to Germany in early childhood. As a candidate for the office of the mayor of Berlin in 2014, he participated in a TV debate. After the debate, Saleh’s German skills were judged as inadequate because of many grammatical mistakes. However, a journalistic analysis [1] demonstrated that Saleh’s use of German grammar was in fact better (one mistake every 140 words) than the use of grammar of a native German TV host in the debate (one mistake every 34 words). It is possible that Saleh’s migrant background and non-standard pronunciation (e.g., using the palato-alveolar [ʃ] instead of the palatal [ç]) triggered this illusion of ungrammaticality.

Correspondingly, empirical investigations revealed that foreign accents modulate speech perception. Accented speech is more challenging to

understand than native productions [2]. It has been experimentally linked to a general processing slow-down [3, 4] and increased processing cost [5]. Generally, non-native productions deviate from native utterances on a number of linguistic properties. The degree of accentedness has been shown to strongly depend on segmental errors [6, 7], and distorted suprasegmental features, such as intonation, timing, speech rate, and loudness [6, 8, 9]. Additionally, L2 speakers often use grammatical structures deviating from native language usage. The presence of grammatical errors could increase the degree of perceived accentedness, but results from the few available studies are mixed. Munro & Derwing [7] found a correlation between the number of grammatical errors spontaneously produced by accented or native speakers and the accent scores. However, when Asano & Weber [10] manipulated the grammaticality of German sentences produced by L1 and L2 speakers, participants rated speech containing grammatical errors as more foreign only when it was produced by L1 but not by L2 speakers. Furthermore, evidence from EEG research suggests that the grammaticality of utterances is evaluated differently for native and foreign-accented speech. In an ERP study by Hanulíková et al. [11], grammatical errors in sentences spoken with a foreign accent did not lead to a P600 effect, typically evoked by the same errors in sentences spoken by native speakers. Because the P600 component is commonly associated with structural reanalysis and repair, the authors interpreted their findings as an effect of listeners expecting incorrect utterances from foreign-accented speakers and consequently reducing attempts to repair such mistakes. The same expectation of ungrammaticality that results in reduced sensitivity to errors might also prime listeners to perceive illusory errors in well-formed utterances, resulting in a so-called ‘grammatical tinnitus’.

Research has also shown that a foreign accent can trigger specific biases in listeners reflected, for example, in decreased ratings of traits like status, solidarity, or dynamism for accented speakers [12]. Crucially, such biases and expectations can also exert a top-down influence on perceptual processes, as reported for the phoneme categorization [13] and lexical recognition [14] tasks.

The discussion of the Raed Saleh's candidacy, as well as the research findings listed above motivated our first hypothesis that the number of incorrect responses in a grammaticality judgment task for grammatical sentences (both well-formed and phonologically anomalous) will be higher in the foreign-accented speech condition than in the native-accented speech condition. We also expect that the foreign accents will lead to a greater increase in the number of incorrect responses in sentences with phonological anomalies in comparison to well-formed sentences.

## 2. EXPERIMENTS

### 2.1. Experimental design

To test our hypotheses, we conducted two mirror experiments taking place in Germany and Poland. Participants listened to native and foreign-accented sentences further divided into well-formed, phonologically anomalous, and ungrammatical utterances. Sentences from the foreign accent condition were recorded by L2 speakers and were marked mostly by a subtle accent characterized by suprasegmental variation. Both experiments combined the ERP technique of recording brain potentials with the behavioral task of grammaticality judgments. As part of the task, participants had to decide whether the auditorily presented sentence was grammatically correct. Only the behavioral results of the study are reported below, and we limit our discussion to the well-formed and phonologically anomalous conditions essential for answering the hypotheses concerning the 'grammatical tinnitus'. The ungrammatical condition served to provide clear cases of ungrammaticality for the judgment task.

#### 2.1.1. Materials

For each experiment, we selected 180 unique experimental items of comparable length following the same syntactic structure. For every item, three sentence versions were created: well-formed, phonological substitution, and grammatical error. This resulted in 540 experimental sentences. The critical middle region started with a preposition followed by a possessive determiner and a two or three syllabic noun. The three versions of each item differed only in the critical region, see Examples (1) and (2) below. For German and Polish audio examples please visit the project's OSF repository ([https://osf.io/4w2ct/?view\\_only=d005b1e78dba4d1a8c7b07eebff7e56f](https://osf.io/4w2ct/?view_only=d005b1e78dba4d1a8c7b07eebff7e56f)). In the phonologically anomalous condition, three native vowels were replaced in an equal number of sentences by sounds

easier to pronounce by L2 learners. Note that Polish learners of L2 German typically struggle with the German lax-tense contrast, e.g., [ɪ]. For German data, we decided to use opposite vowel substitutions. Additionally, we have chosen substitutions between [ʏ] and [i], since [ʏ] does not belong to the Polish phoneme system and [i] is not part of the German vowel inventory. In the German experiment, the substitutions were: [e]→[ɛ], [o]→[ɔ], and [ʏ]→[i]. In the Polish experiment, the substitutions were [ɛ]→[e], [ɔ]→[o], and [i]→[y]. The phonological substitutions always appeared in the stressed, penultimate syllable. Grammatical errors consisted of L2-typical gender mismatches between the possessive determiner and the noun, see e.g., [16].

#### (1) Example: German sentence triple

*Lena befragt die Lehrerin ...*  
Lena consults the teacher

##### (i) well-formed

*...zu ihrem Fehler in der Klausur.*  
about her<sub>[masc]</sub> mistake<sub>[masc]</sub> in the exam

##### (ii) phonological substitution

*...zu ihrem F[ɛ]ler in der Klausur.*  
about her<sub>[masc]</sub> mistake<sub>[masc]</sub> in the exam

##### (iii) grammatical error

*...zu ihrer Fehler in der Klausur.*  
about her<sub>[fem]</sub> mistake<sub>[masc]</sub> in the exam

#### (2) Example: Polish sentence triple

*Nina robi porządek...*  
Nina makes order

##### (i) well-formed

*...w swojej kwaterze przed przyjazdem gości.*  
in her<sub>[fem]</sub> flat<sub>[fem]</sub> before arrival guests

##### (ii) phonological substitution

*...w swojej kwat[e:]rze przed przyjazdem gości.*  
in her<sub>[fem]</sub> flat<sub>[fem]</sub> before arrival guests

##### (iii) grammatical error

*...w swoim kwaterze przed przyjazdem gości.*  
in her<sub>[masc]</sub> flat<sub>[fem]</sub> before arrival guests

Digital auditory recordings of the materials were created. For each experiment, half of the sentences were recorded by native speakers and the other half by L2 speakers with a non-native accent. The same four phonologically trained male speakers recorded sentences for both the Polish and German experiment: two native speakers of Polish who were also L2 speakers of German and two native speakers of German who were also L2 speakers of Polish. The experimenters provided instructions for the speakers how to produce segmental violations in the critical

word, avoiding segmental changes in other parts of the sentence. Additionally, the recordings were inspected by two phonetically trained linguists for categorical changes and their accent strength. If a file contained any other categorical violations than the intended one, it was re-recorded. Thus, materials belonged to six conditions: 3 related to Sentence Type (well-formed, phonological substitution, grammatical error) and 2 related to Sentence Accent (native, foreign). The different versions of experimental items were distributed across three lists such that each participant listened to only one version of a given item. 180 experimental sentences (30 per condition) appeared on one list. Additionally, 72 fillers (36 native-accented, 36 foreign-accented) were recorded by the same speakers and added to each list. This resulted in a total number of 252 sentences per list.

### 2.1.2. Procedure

The experiments were conducted in EEG laboratories in Germany and Poland. Participants were seated in front of a computer screen. The experimental session started with written instructions appearing on the screen. Participants were asked to listen to the recordings of sentences while looking at the screen and to respond to questions following the sentences. Once the recording stopped, the question ‘Is the sentence grammatically correct?’ was displayed in the respective language. The participants’ task was to respond by pressing one key for ‘Yes’ or another for ‘No’ as fast and accurately as possible. Their responses were recorded. A comprehension question appeared after the grammaticality question in one third of trials. A short practice session preceded the experiment proper.

### 2.1.3. Participants

33 native speakers of German (17 women, 16 men) aged between 21 and 35 years ( $M=26$ ,  $SD=3.8$ ) participated in the German experiment. They were mostly students of Phillips-Universität Marburg, except for 8 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.

## 2.2. Statistics

All statistical analyses were conducted with R (version: 4.2.0, [17]). A binomial logistic regression

model (package ‘lme4’, [18]) was fitted with JUDGMENT ACCURACY [correct, incorrect] as the dependent variable. Judgment accuracy values were derived from ‘Yes’ or ‘No’ answers to the question ‘Is the sentence grammatically correct?’. For grammatical (well-formed and phonologically anomalous) sentences, ‘No’ constituted an incorrect answer. For ungrammatical sentences, it would constitute a correct answer. We fitted separate models for the German and Polish experiment. In both models, the fixed factors were ERROR TYPE [well-formed, phonological substitution, grammatical error], ACCENT TYPE [native, foreign], and their interactions. We also included PARTICIPANT and SENTENCE TRIPLE as random intercepts and ERROR TYPE, ACCENT TYPE and their interaction as slopes. Because of convergence issues, we removed the interactions between ERROR TYPE and ACCENT TYPE from all random structures. For multiple comparisons we used the ‘emmeans’ function [19].

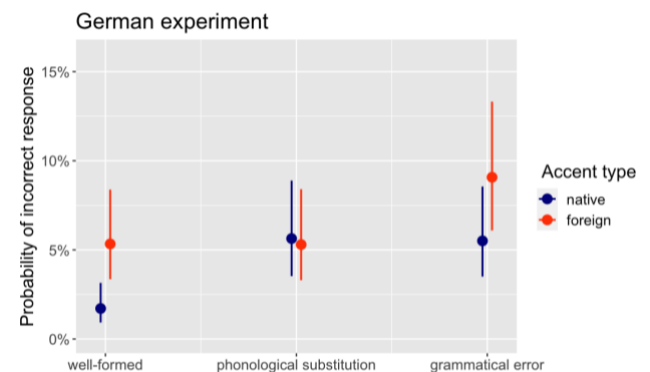
For the statistical analysis of German data, 5917 data points were submitted. For Polish, 5369 data points were analysed.

## 3. RESULTS

### 3.1. German experiment

The analysis of the German experiment revealed a significant interaction for ERROR TYPE and ACCENT TYPE between well-formed sentences and phonologically anomalous sentences ( $z=-3.49$ ,  $p<.001$ ).

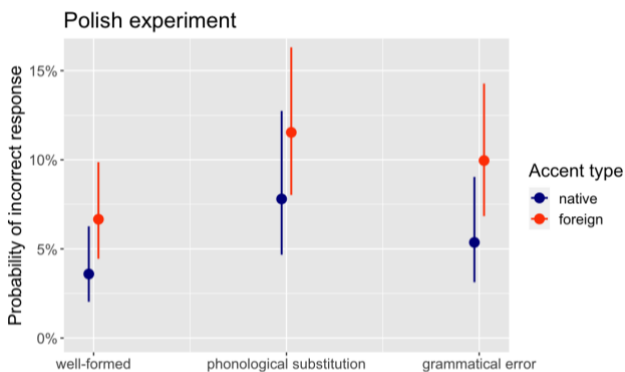
The results indicate that sentences with no grammatical or phonological error (well-formed), spoken with a foreign accent were more likely to be (incorrectly) judged as ungrammatical than well-formed sentences spoken with a native accent ( $z=-3.3$ ,  $p<.05$ ), see Figure 1. A comparison of sentences with a phonological substitution revealed no significant difference between the native and foreign accent condition.



**Figure 1:** Probability of incorrect response for Error Type and Accent Type, German experiment

### 3.2. Polish experiment

The data analysis of the Polish experiment revealed a simple effect of ACCENT TYPE in the well-formed condition, with foreign-accented sentences being more likely to be judged as ungrammatical than native-accented sentences ( $z=3.25, p<.01$ ), see Figure 2. Likewise, sentences containing phonological anomalies were more likely to be judged as ungrammatical when being foreign-accented compared to native-accented sentences ( $z=2.33, p<.05$ ).



**Figure 2:** Probability of incorrect response for Error Type and Accent Type, Polish experiment

However, the statistical modelling indicated no significant interaction between ERROR TYPE and ACCENT TYPE. As demonstrated in Figure 2, the overall probability of incorrect responses increased for both the native and the foreign accent when sentences contained phonological anomalies.

### 4. CONCLUSIONS

The present study investigated whether listeners exposed to foreign-accented speech experience a form of perceptual illusion where non-existent grammatical errors are heard in grammatically correct speech. We predicted that the number of incorrect grammaticality judgments in the well-formed and phonologically anomalous condition would be higher for foreign-accented, compared to native-accented speech. Additionally, we predicted the number of incorrect responses to increase for sentences with vowel substitutions in the critical phrase.

The data from both languages suggest that listeners were more likely to judge well-formed utterances with no segmental substitutions as ungrammatical when they were uttered by foreign-accented speakers compared to fully native-sounding speakers. This effect can be interpreted as the ‘grammatical tinnitus’, a form of (un)grammaticality illusion. Various accounts of the mechanism underlying this phenomenon are possible. First, the results are consistent with the hypothesis that

exposure to non-native accents induces a perception of non-existent grammatical errors in well-formed sentences. In other words, the locus of the effect would be at the online perception and comprehension stage. Second, the sentence as a whole can be judged as ungrammatical without the participant actually hearing any ‘phantom’ errors during speech perception.

Results also revealed unexpected differences between the experiments. In Polish, the overall probability of judging utterances as ungrammatical increased both in well-formed sentences and sentences containing phonological substitutions, as hypothesized. Surprisingly, the German data revealed no difference between the native and foreign-accented sentences in the phonologically anomalous condition, suggesting that Polish listeners were affected more by phonological anomalies than German listeners. We propose three explanations for the differences between the languages. Firstly, Germans might have a greater experience with segmental variation because of their exposure to various dialects, sociolects, and foreign accent variants. Due to their familiarity with segmental variation, they might be less sensible to phonological substitutions, leading to less confusion about the grammatical correctness of an utterance. Secondly, there might be differences between the degree of familiarity with the vice-versa accent. Previous ERP studies [20, 21] indicated that more familiarity with an accent leads to more native-like processing of language errors. Due to the awareness of the segmental patterns of a particular accent, listeners might be less prone to judge segmental violations as grammatical errors, resulting in no extra ‘grammatical tinnitus’. However, a questionnaire filled by the participants showed little difference in self-reported—exposure to the non-native accents between both groups. Thirdly, the perceived strength of segmental violations might have diverged between the German and Polish experiments. While for two of three segmental violations the change was from tense to lax in the German experiment, it was the other way around for the Polish experiment.

In summary, the present investigation demonstrates that a foreign accent affects the way listeners perceive the grammaticality of utterances. The ‘grammatical tinnitus’ can result in foreign speakers being perceived as less (linguistically) competent, as illustrated by the example of Raed Saleh.

## 5. ACKNOWLEDGMENT

This work was supported by Deutsch-Polnische Wissenschaftsstiftung within the project ‘Grammatical tinnitus and its role in the perception of foreign language accent. A comparison of German and Polish’ (grant number: 2020-07). We are grateful to Ulrike Domahs for giving Sarah Wesolek the opportunity to run the experiment with German speakers at Marburg University. Finally, we also thank the Leibniz-Society.

## 5. REFERENCES

- [1] Heiser, Sebastian. 2014. Berlins Bürgermeisterkandidat Saleh: Ein dubioses Hörproblem. *TAZ: Gesellschaft/Medien*. [Online]. <https://taz.de/Berlins-Buergermeisterkandidat-Saleh/!5034466/> [Accessed: January 6, 2023].
- [2] Bent, T., Bradlow, A. 2003. The Interlanguage Speech Intelligibility Benefit. *The Journal of the Acoustical Society of America*, 114, 1600–1610.
- [3] 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.
- [4] 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.
- [5] Van Engen, K. J., Peelle, J. E. 2014. Listening effort and accented speech. *Frontiers in Human Neuroscience*, vol. 8, 1–4.
- [6] Magen, H. S. 1998. The perception of foreign-accented speech. *Journal of Phonetics*, 26 (4), 381–400.
- [7] 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.
- [8] Munro, M. J. 1995. Nonsegmental Factors in Foreign Accent: Ratings of Filtered Speech. *Studies in Second Language Acquisition*, 17 (1), 17–34.
- [9] de Mareüil, P. B., Vieru-Dimulescu, B. 2006. The Contribution of Prosody to the Perception of Foreign Accent. *Phonetica*, 63 (4), 247–267.
- [10] 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.
- [11] Hanulíková, A., van Alphen, P. M., van Goch, M. M., Weber, A. 2012. When one person’s mistake is another’s standard usage: the effect of foreign accent on syntactic processing. *J Cogn Neurosci*, 24 (4), 878–887.
- [12] 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.
- [13] Jannedy, S., Weirich, M. 2014. Sound change in an urban setting: Category instability of the palatal fricative in Berlin. *Laboratory Phonology*, 5, 91–122.
- [14] Ingvalson, E. M., Lansford, K. L., Federova, V., Fernandez, G. 2017. Listeners’ attitudes toward accented talkers uniquely predicts accented speech perception. *The Journal of the Acoustical Society of America*, 141 (3), EL234–EL238.
- [15] 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.
- [16] 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.
- [17] 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/>.
- [18] Bates, D., Mächler, M., Bolker, B., & Walker, S. 2015. Fitting linear mixed-effects models Using lme4. *Journal of Statistical Software*, 67 (1).
- [19] Lenth, R. 2022. emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.8.3, <https://CRAN.R-project.org/package=emmeans>.
- [20] Grey, S., van Hell, J. G. 2017. Foreign-accented speaker identity affects neural correlates of language comprehension. *Journal of Neurolinguistics*, 42, 93–108.
- [21] Holt, R., Kung, C., Demuth, K. 2018. Listener characteristics modulate the semantic processing of native vs. foreign-accented speech. *PLoS ONE*, 13 (12), 1–24.