# PERCEPTUAL ASSIMILATION AND FREE CLASSIFICATION OF GERMAN VOWELS BY AMERICAN ENGLISH LISTENERS

Danielle Daidone, Franziska Kruger, and Ryan Lidster

Indiana University
ddaidone@indiana.edu, fkruger@indiana.edu, rflidste@umail.iu.edu

## ABSTRACT

In order to investigate listeners' perceptual patterns regarding segments from an unfamiliar language, researchers commonly employ perceptual assimilation tasks. These tasks investigate the perceptual similarity of non-native sounds to L1 sounds, but do not provide information on the perceived similarity among non-native sounds. In order to examine how German vowels are perceived by naïve American English listeners, both in their similarity to L1 vowels and to each other, this study employs a perceptual assimilation task and a free classification task, a tool previously used for investigating perceived (dis-)similarities of stimuli.

The perceptual assimilation results largely replicate previous findings on American English assimilation of German vowels; however, the free classification results suggest that assimilation patterns are not reliable indicators of the German vowels' perceptual similarity to each other.

These results indicate that free classification offers an efficient means of gathering corroborating data that can be used to enhance cross-linguistic perception research.

**Keywords**: perceptual assimilation, free classification, German, front rounded vowels

## 1. INTRODUCTION

According to predominant models of speech perception, the first language (L1) strongly influences the perception of foreign sounds [2, 12]. Difficulty in perceiving a particular non-native sound contrast is said to be determined by the perceptual similarity of the foreign sounds to L1 phonemes.

In order to predict which contrasts will be difficult for an unexperienced listener or a learner of a foreign language, researchers have often employed perceptual assimilation (PA) tasks to determine perceptual similarity patterns, e.g. [8, 13]. In PA tasks, participants listen to segments of an unfamiliar language, choose which L1 segment is most similar, and rate how good that segment is as an example of that L1 category. PA tasks provide information on how non-native sounds relate to L1 categories, and researchers have used these results to make inferences about how (dis-)similar non-native sounds are to each other [13]. In particular, two non-native sounds that are assimilated to the same L1 sound are assumed to be difficult to discriminate, while foreign sounds perceived as distinct L1 sounds will be easier to distinguish [2]. However, PA does not directly investigate the perceptual distances of non-native sounds to each other, and categorization patterns are often not straightforward, complicating their interpretation.

For example, Strange et al. [13] found that North German /ʏ/ and /ø/ were both categorized as multiple American English (AE) vowels, including AE /ʊ/ 56% and 28% of the time, respectively, and both were rated as poor examples of all of the vowels they were categorized as. From these results, it is challenging to predict how similar AE listeners perceive /ø/ and /ʏ/ to be. Since they were considered poor examples of English vowels, this might indicate that they both differed in the same way from English vowels, especially since both were sometimes assimilated to /ʊ/, in which case, they would be perceptually similar and discrimination would be difficult. Alternatively, it is possible that these vowels were judged as poor examples of English vowels for different reasons, which would suggest that these sounds would be easily discriminated. Only an investigation of the perceptual distances among the sounds of the non-native language can speak to these types of situations.

In laboratory phonology, perceptual distances among non-native sounds have been investigated using a variety of tasks, including AX, Oddball, AXB/ABX, Oddity, and sequence recall tasks [14, 7]. All these tasks, however, require large numbers of trials per contrast. In the case of German, using any of these tasks to examine the similarity of all 14 vowel phonemes to each other would require a prohibitively large number of trials.

Auditory free classification tasks have also been employed to study the perceptual similarity of stimuli, but so far have only been used to examine the similarity of regional dialects [5, 6], of various foreign accents [1], and of different languages to each other [4]. In a free classification task, participants are asked to create groups of similar-sounding items with no category labels imposed by the experimenter. Avoiding pre-specified categories greatly reduces

response biases due to category labels [6] and allows participants to group the stimuli according to dimensions they find relevant, which may be a dimension unexpected by the researcher and/or not easily explained to participants such as vowel reduction or roundedness [4]. Additionally, completing the task does not require knowledge of the orthography of the target language or any linguistic terminology. Furthermore, unlike other perceptual similarity judgment tasks in which listeners complete numerous trials that each present two sounds to compare [10], free classification tasks present all the stimuli for comparison at once, and thus take far less time to complete. To date, however, free classification tasks have not been used to examine the perceived similarity of segments.

This study examines whether free classification tasks yield interpretable results regarding the perceptual similarity of foreign segments, and compares the perceptual distances between stimuli observed in free classification to L1 assimilation patterns obtained from a PA task. By comparing the pattern of perceptual similarity among foreign segments in two tasks—with and without category labels and reference to the L1—the goal of the study is to examine whether PA tasks provide sufficiently reliable data to determine the perceived similarity of non-native sounds to each other.

Our study examines German vowels as perceived by American English (AE) listeners. In previous research using a PA task [13], AE listeners assimilated German front and back rounded vowels to AE back rounded vowels, whereas front unrounded vowels were assimilated to AE front vowels. We predict that perceptual distances obtained in the free classification task will reflect those patterns.

## 2. METHOD

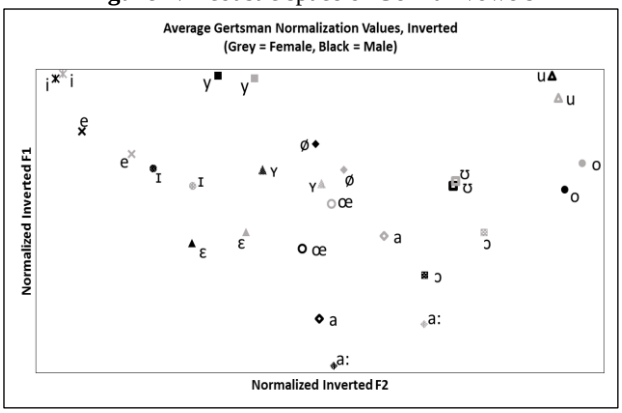
### 2.1. Participants

Thirty-five AE speakers participated in the experiment. Results from six of the participants were excluded: five had exposure to languages with front rounded vowels, and one did not correctly categorize L1 vowels in the training phase. The remaining 29 participants (age range 20-25) had no prior exposure to German or any other language with front rounded vowels and were all classified as naïve listeners. All participants passed a bilateral hearing screening and none reported any prior hearing or speech problems.

### 2.2. Stimulus material and vowel analysis

Four native speakers of German (2 female from Brandenburg and Saxony, 2 male from Rhineland-Palatinate, age ranges 31-45) read a list of monosyllabic CCVC stimuli produced in the sentence *Ich sage…* ("I say…"). Stimuli incorporated the German vowels /i/, /ɪ/, /y/, /ʏ/, /u/, /ʊ/, /e/, /ɛ/, /o/, /ɔ/, /ø/, /œ/, /a:/, and /a/ in alveolar (/ʃtVt/) and velar context (/skVk/). Stimuli for the familiarization phase of the PA task were English monosyllabic words and non-words recorded by two native speakers of AE (1 female (age 26), 1 male (age 30)). The recordings of the German speakers were analyzed in Praat [3]. Vowel boundaries were marked at a zero-crossing at the beginning and end of each vowel and the values for F0, F1, F2 and F3 were extracted for the tokens of each speaker at 25%, 50% and 75% of the total vowel duration. Following Flynn [9], formants were normalized with Gerstman's formula. The average normalized F1 and F2 values at the vowel midpoint for the male and female German speakers' vowel productions used in the free classification task (including both /skVk/ and /ʃtVt/ contexts) are shown in Figure 1.

**Figure 1**: Acoustic space of German vowels




### 2.3. Procedure

Participants completed a hearing screening, free classification task, background questionnaire, and PA task, all administered via headphones in that order. Free classification was administered via a PowerPoint presentation consisting of two slides, each containing a 16x16 square grid on the left-hand side and the sound files in randomly numbered squares on the right-hand side. Each slide presented 28 items (14 vowels x 1 context x 2 speakers), and order of slides was counterbalanced across participants.

Participants were instructed to drag the squares onto the grid to make groups of similar-sounding vowels. They could listen to each item and rearrange groups as often as they liked. Participants were told to make groups of at least two sounds but no upper limit was imposed on group size. They were asked to pay attention to the quality of the vowels while ignoring speaker sex and were not given any category labels. Participants took approximately 10-20 minutes to complete the task.

The PA task was completed in Praat. Each trial consisted of the auditory stimulus presentation, response selection and similarity rating. Thirteen response choices labeled with keywords (*he* for /i/, *hid* /ɪ/, *hey* /eɪ/, *head* /ɛ/, *had* /æ/, *ha* /ɑ/, *haw* /ɔ/, *huh* /ʌ/, *who* /u/, *hood* /ʊ/, *hoe* /oʊ/, *her* /ɝ/, and *other*) were presented as buttons on-screen. Listeners were asked to select the keyword with the vowel most representative for the respective auditory stimulus and to rate its similarity on a 6-point scale (1 being very *different* to the AE category, 6 being *similar*). They were instructed to treat /ɝ/ as a vowel. They were also told to use the button labeled 'other' if they could not assign a stimulus to any of the AE vowel categories. Trials were self-paced. First, listeners completed a training phase with 23 English words and non-words to emphasize the focus on the similarity of vowels and minimize the likelihood that participants would rely on consonantal context. Following the training phase, participants were asked to read all key words aloud to verify that they had the appropriate vowels in mind when completing the tasks. The experimental phase consisted of two blocks, each consisting of 112 trials (14 vowels x 2 contexts x 4 speakers) for 224 trials in total. Presentation of the stimuli was randomized within each block. Participants took about 30-35 minutes to complete this task.

### 3. ANALYSIS AND RESULTS

For both tasks, small but significant consonantal context effects were observed. However, since they were very consistent across the PA task and free classification task, the present analysis uses the combined results of both consonantal contexts.

For the PA task, listeners' choices of L1 category for each German vowel were calculated in percent response rates. Response rates lower than 2.5% were disregarded. The results are summarized in Table 1. Nine vowels (/y/, /i/, /ɪ/, /ɛ/, /ø/, /œ/, /a:/, /ɔ/, /u/) were classified in one category more than 50% of the time. The remaining five vowels (/ʏ/, /e/, /a/, /o/, /ʊ/) were more ambiguous.

German front unrounded vowels were assimilated to AE front unrounded vowel categories. In contrast, the German front rounded vowels (with the exception of 3.5% of the tokens of /œ/ classified as /ɛ/) were not classified as front vowels at a rate of 2.5% or higher. Rather, German tense front rounded vowels were most often assimilated to AE back vowel categories (/ø/ to AE /ʊ/; /y/ to AE /u/ and /ʊ/), and lax front rounded vowels were assimilated to either lax back or central AE vowels (/ʏ/ to AE /ʊ/ and /ʌ/; /œ/ to AE /ʌ/). For all front rounded vowels, /ɝ/ was an infrequent response. In contrast, "other" was chosen very rarely (<0.3% of total trials).

**Table 1**: Perceptual assimilation response rates (%)

| Keyword | | he | hid | hey | head | had | huh | her | ha | haw | hoe | hood | who | other |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AE Category | | i | ɪ | eɪ | ɛ | æ | ʌ | ɝ | ɑ | ɔ | oʊ | ʊ | u | |
| German Vowel | y | | | | | | 3.5 | | | | | 33.8 | **59.5** | |
| | ʏ | | | | | | 30.7 | 3.7 | | | | **49.8** | 8.0 | |
| | i | **82.9** | 7.8 | 4.8 | 3.5 | | | | | | | | | |
| | ɪ | 3.5 | **67.7** | 13.9 | 13.4 | | | | | | | | | |
| | e | 19.3 | 14.8 | **44.7** | 14.8 | 3.3 | | | | | | | | |
| | ɛ | | 11.1 | 19.9 | **59.6** | 4.8 | | | | | | | | |
| | ø | | | | | | 12.3 | 4.3 | | | | **60.4** | 19.3 | |
| | œ | | | | 3.5 | | **55.9** | 3.7 | 5.5 | 3.7 | | 18.8 | 2.8 | |
| | a | | | | | 7.6 | 32.2 | | **38.5** | 15.2 | | | | |
| | a: | | | | | 10.4 | 4.1 | | **52.6** | 29.8 | | | | |
| | o | | | | | | 2.8 | | | **39.0** | 24.1 | 17.1 | 11.7 | |
| | ɔ | | | | | 4.1 | 18.7 | | 15.2 | **51.1** | 2.8 | 5.7 | | |
| | u | | | | | | | | | | | 31.0 | **64.0** | |
| | ʊ | | | | | | 28.2 | | 3.1 | 16.2 | 9.4 | **31.9** | 7.2 | |

Category goodness ratings were, in general, very difficult to interpret. For example, German /y/ was categorized as AE /u/ more often than as AE /ʊ/ (59.5% to 33.8%). However, goodness ratings for /y/ were higher for AE /ʊ/ (4.5/6) than for AE /u/ (4.11/6). The results for German /u/ patterned similarly (/ʊ/: 31.0% (4.77/6); /u/: 64.0% (4.49/6)). To aid in interpreting complicated PA results, Guion et al. [11] proposed a "fit index" calculated as the product of the proportion of category selection and the average goodness ratings. Table 2 shows the fit index values, calculated for all categories with a response rate of >20%. Fit index values indicated that AE /u/ was a better overall match for German /y/ and /u/ than AE /ʊ/.

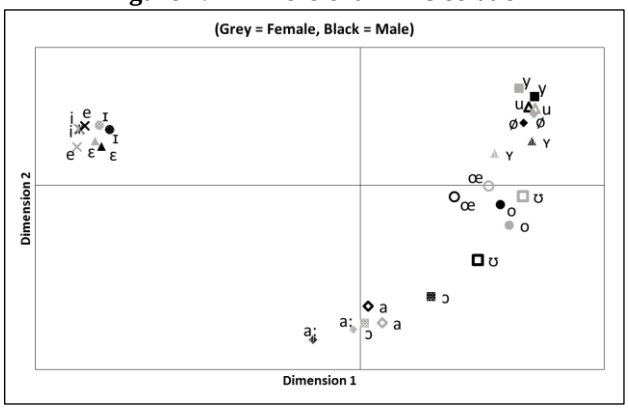
**Table 2**: PA Fit Indices

| Keyword | | he | hid | hey | head | had | huh | her | ha | haw | hoe | hood | who | other |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AE Category | | i | ɪ | eɪ | ɛ | æ | ʌ | ɝ | ɑ | ɔ | oʊ | ʊ | u | |
| German Vowel | y | | | | | | | | | | | 1.52 | 2.44 | |
| | ʏ | | | | | | 1.13 | | | | | 1.93 | | |
| | i | 4.03 | | | | | | | | | | | | |
| | ɪ | | 3.03 | | | | | | | | | | | |
| | e | | | 2.03 | | | | | | | | | | |
| | ɛ | | | | 2.64 | | | | | | | | | |
| | ø | | | | | | | | | | | 2.58 | | |
| | œ | | | | | | 2.21 | | | | | | | |
| | a | | | | | | 1.31 | | 1.72 | | | | | |
| | a: | | | | | | | | 2.48 | 1.33 | | | | |
| | o | | | | | | | | | 1.51 | 1.04 | | | |
| | ɔ | | | | | | | | | 2.21 | | | | |
| | u | | | | | | | | | | | 1.48 | 2.88 | |
| | ʊ | | | | | | 1.09 | | | | | 1.26 | | |

For the free classification task, tokens were coded dichotomously as either grouped or not grouped with each other and summed across subjects to create dissimilarity matrices for each consonant context which were analyzed using SPSS 22's ALSCAL function for multidimensional scaling (MDS) with a convergence criterion of 0.001 [5]. MDS attempts to position the tokens in an arrangement that recreates the differences in the dissimilarity matrix. After examining solutions ranging from 1 to 5 dimensions, a two-dimensional solution was deemed optimal in that it provided adequate data fit ($R^2 > 0.88$, stress =

.174) and was located below the elbow of the stress plot. The two-dimensional MDS solution is presented in Figure 2.

**Figure 2**: 2-Dimensional MDS solution




In order to interpret the two dimensions, we calculated correlations between the dimension scores of the tokens and their respective formant measurements, pitch, intensity, duration, and change in formant values or pitch over the vowel duration. Additionally, speaker sex and vowel roundedness were entered as dichotomous variables. The results, shown in Table 3, indicate that, for the first dimension, although F2 and F3 were significantly correlated with dimension scores, roundedness of the vowel was the strongest predictor of perceptual distance in that respect. For Dimension 2, F1 values showed the highest correlation. F0, intensity, duration, and speaker sex were not strongly related to either dimension, nor was any measurement of change in formant values or pitch.

**Table 3**: Correlations with total MDS Solution

| | F0 | F1 | F2 | F3 | Dur | dB | Sex | Rounded |
|---|---|---|---|---|---|---|---|---|
| Dim1 | 0.01 | -0.04 | -0.73 | -0.69 | 0.01 | 0.13 | 0.00 | **0.88** |
| Dim2 | -0.09 | **0.81** | -0.42 | -0.39 | 0.16 | -0.27 | -0.01 | -0.25 |

## 4. DISCUSSION

Our PA results, in general, replicate previous findings on AE assimilation of German vowels [13]. German tense front rounded vowels were mostly assimilated to English back vowel categories, while lax front rounded vowels were primarily split between AE /ʊ/ and /ʌ/. Our listeners, however, were much less uniform in their classifications of vowels than the listeners in Strange et al.'s [13] experiment. While the exact formant values of our German vowel productions differed slightly, it seems more likely that the differences can be attributed to our recruitment of participants who had not had training in phonetics or general linguistics. Multiple checks during the training phase indicated that our participants understood the task, but it is possible that their lack of metalinguistic knowledge lowered the confidence and consistency of their classifications of perceptually ambiguous tokens. Future research in this paradigm should consider the potential effect of linguistic training when selecting participants.

While the PA results show how German vowels were assimilated to AE categories, the free classification results strongly suggest that assimilation patterns do not clearly indicate the German vowels' perceptual similarity to each other. For example, German /ʏ/, /ʊ/, and /a/ were all assimilated to AE /ʌ/ approximately 30% of the time (with non-significantly different goodness ratings). Guion et al. [11] suggested using the fit index to rank similarity of multiple non-native sounds classified as the same L1 category. However, the three vowels obtained fit index values of 1.13, 1.09, and 1.31 for /ʌ/, which might be interpreted as a sign that /ʏ/, /ʊ/, and /a/ were perceived as similar to each other.

The MDS results suggest that this was not the case. While /ʏ/ was indeed commonly grouped together with /ʊ/, both vowels were perceived as very distant from /a/. Separately, the assimilation patterns for /u/ and /ʏ/ are dissimilar, and yet listeners very frequently grouped the two vowels together in the free classification task, indicating that they perceived them as similar to each other, even if individually they would be identified as corresponding to different L1 categories. Thus, the free classification results indicate that PA alone is not sufficient to predict the perceived similarity of foreign sounds to each other.

Free classification was also useful in providing an easily interpretable representation of listeners' perceptual space. Indeed, the MDS solution for the perceived distribution of non-native sounds is similar to their distribution in acoustic space, only with the front rounded vowels perceived as more similar to back vowels. Analyzing the dimension scores revealed that vowel height (F1) was a salient dimension, but that roundedness was a stronger predictor than backness (F2) for predicting AE listeners' perceptual patterns regarding German vowels.

In sum, our study suggests that, in addition to the domains in which free classification had been applied previously, the task can provide useful data on the perceived similarity of non-native segments. Free classification does this in a relatively short period of time without the need to impose the number or nature of categories, and should thus be considered a viable option for cross-linguistic perception research.

## 5. ACKNOWLEDGEMENTS

We thank Ken de Jong, Isabelle Darcy, Eriko Atagi, and the IU L2 Psycholinguistics Lab for their valuable feedback.

## 6. REFERENCES


[1] Atagi, E., Bent, T. 2013. Auditory free classification of nonnative speech. *Journal of Phonetics* 41, 509-519.

[2] Best, C. T. 1995. A direct-realist view of cross-language speech perception. In: W. Strange (ed), *Speech perception and linguistic experience: Issues in cross-language speech research.* York Timonium, MD: York Press, 171-206.

[3] Boersma, P., Weenink, D. 2015. Praat: doing phonetics by computer. Version 5.4.08 http://www.praat.org/.

[4] Bradlow, A. R., Clopper, C., Smiljanic, R., Walter, M. A. 2010. A perceptual phonetic similarity space for languages: Evidence from five native language listener groups. *Speech Communication* 52, 930–942.

[5] Clopper, C. G. 2008. Auditory free classification: Methods and analysis. *Behavior Research Methods* 40, 575-581.

[6] Clopper, C. G., Pisoni, D. B. 2007. Free classification of regional dialects of American English. *Journal of Phonetics* 35, 421-438.

[7] Dupoux, E., Sebastián-Gallés, N., Navarrete, E., Peperkamp, S. 2008. Persistent stress 'deafness': The case of French learners of Spanish. *Cognition* 106, 682-706.

[8] Escudero, P., Vasiliev, P. 2011. Cross-language acoustic similarity predicts perceptual assimilation of Canadian English and Canadian French vowels. *J. Acoust. Soc. Am.* 130, EL277-EL283.

[9] Flynn, N. 2011. Comparing vowel formant normalisation procedure. *York Papers in Linguistics Series 2* 11, 1-28.

[10] Gandour, J. 1984. Tone dissimilarity judgments by Chinese listeners. *Journal of Chinese Linguistics* 12, 235-261.

[11] Guion, S. G., Flege, J. E., Akahane-Yamada, R., Pruitt, J. C. 2000. An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *J. Acoust. Soc. Am.* 107, 2711-2724.

[12] Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., Nelson, T. 2008. Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences* 363, 979-1000.

[13] Strange, W., Bohn, O.-S., Nishi, K., Trent, S.A. 2005. Contextual variation in the acoustic and perceptual similarity of North German and American English vowels. *J. Acoust. Soc. Am.* 118, 1751-1762.

[14] Strange, W., Shafer, V. L. 2008. Speech perception in second language learners: The re-education of selective perception. In: Handsen Edwards, J. G., Zampini, M. L. (eds), *Phonology and Second Language Acquisition*. Philadelphia: John Benjamin, 153-191.