

RECIPROCAL MEASURES OF PERCEPTUAL SIMILARITY

Juli Cebrian

Universitat Autònoma de Barcelona
juli.cebrian@uab.cat

ABSTRACT

This paper reports the results of two experiments aimed at assessing the perceived similarity between a set of English and Catalan vowels from the perspective of native speakers of both languages. Two groups of 27 listeners, a group of native English speakers and a group of native Catalan speakers, used a 9-point scale to rate the degree of similarity between two stimuli. Crucial stimuli consisted of English-Catalan vowel pairs. The experiments also included same-category pairs from a single language.

The results of the experiments showed that some mixed-language vowel pairs obtained similarity ratings that fell within the range of same-category single-language pairs, revealing a perceptual overlap for some native and non-native phonetic categories. Further, Catalan listeners provided overall higher dissimilarity ratings than English listeners did, showing no relationship between smaller inventory size and greater within-category variability. The outcomes are discussed in terms of the resulting predictions for second language perception.

Keywords: cross-linguistic perception, vowels, perceptual similarity, second language speech.

1. INTRODUCTION

The evaluation of cross-linguistic similarity is a crucial starting point in research on the acquisition of second language (L2) speech. This is because in order to understand how the L2 speech learner perceives the sounds of a given target language (TL), we must first examine the relationship between the phonetic categories of the first language (L1) and the TL. In fact, influential models of L2 speech acquisition have based their predictions about the likelihood of accurate TL category formation on the degree of similarity between existing L1 categories and the TL categories [1,4]. For instance, Flege [4] claims that in order to establish accurate categories for TL sounds, the L2 learner has to be able to discern differences between TL and L1 sounds. Therefore, evaluating the degree of similarity between native and non-native sounds is a necessary first step for any L2 perception study.

Recent research has advocated for the use of cross-language mapping tasks as the most reliable way of evaluating phonetic similarity [7]. These tasks often involve the identification of non-native sounds in terms of native categories plus goodness of fit ratings (perceptual assimilation task). Flege et al. [5] used a rated dissimilarity task in which listeners were asked to provide dissimilarity judgments on pairs consisting of a native and a non-native vowel. This paper will report the results of a larger study on the cross-linguistic similarity between Catalan and English vowels, focusing on data elicited by means of a rated dissimilarity task.

Most studies on cross-linguistic similarity are unidirectional, that is, explore the similarity between native and non-native vowels from the perspective of speakers of a single language (e.g., [3]). An additional goal of this paper is to explore cross-linguistic similarity by contrasting the perceptual judgments of speakers of two languages on the same native and non-native stimuli. We believe that this reciprocal measure provides a more complete assessment of cross-language similarity.

Specifically, this study examines the degree of similarity between a set of Catalan and English vowels and diphthongs. Catalan has a 7 vowel phonemes (/i e ε a o u/), plus the unstressed vowel schwa, and a number of diphthongs (e.g. /ai, au/). Southern British English has 12 vowels (/i ɪ ε ɜ æ ʌ ɑ ɒ ɔ ʊ u ə/) and 8 diphthongs (e.g., /ai au ei əʊ/). One final goal of the study is to examine if differences in vowel inventory affect the degree of stimulus variability that may be tolerated as within category variability. The vowel dispersion theory (e.g., [6]) claims that the smaller the inventory, the greater the amount of within-category variability. If this has an effect on perception, it is possible that non-native stimuli that are relatively close to native stimuli are perceived as more acceptable, and thus less dissimilar to native categories, by Catalan speakers than by English speakers.

The perceptual similarity between English and Catalan vowels was examined by means of two experiments. In the first experiment, a group of English native speakers rated the perceived dissimilarity between Catalan and English vowels. In the second experiment, a group of Catalan

speakers assessed the perceived similarity between the same set of native and non-native vowels.

2. METHODOLOGY

Crosslinguistic similarity was measured by means of a rated dissimilarity task [5], in which listeners rated the degree of (dis)similarity between two stimuli. Specifically, listeners indicated how similar or dissimilar they perceived the two stimuli in a given pair by means of a 9-point Likert scale. This task was part of a larger study involving English and Catalan native speakers who evaluated a large number of native and non-native vowels and diphthongs (14 English and 11 Catalan monophthongs and diphthongs). Given the large amount of data explored, the native and non-native vowels were divided into two sets (roughly front and back vowels) and each set was tested separately. For the sake of brevity and simplicity, this paper will focus on the results for the front vowels and diphthongs.

Experiment 1 involved English native speakers and Experiment 2 involved Catalan native speakers. All speakers performed a rated dissimilarity task that contained the same set of crucial Catalan-English stimulus pairs. In addition, each experiment included a different set of same-language vowel pairs (i.e., Catalan-Catalan or English-English) for baseline and comparison reasons.

2.1. Stimuli

The stimuli for both experiments consisted of seven English vowels and diphthongs (/i ɪ ε ɜ æ aɪ eɪ/), and six Catalan vowels and diphthongs (/i e ε a ei ai/). The vowels were elicited in b+vowel+t sequences, which is a possible sequence in both languages. The /bVt/ words were elicited from three male native speakers of Standard Southern British English (mean age: 30) and three male native speakers of Eastern Catalan (mean age: 35). The bVt stimuli were embedded in carrier phrases of similar length and structure. Stimuli were recorded in London and Barcelona in a soundproof booth at a sampling rate of 44kHz and digitally stored for further editing. In order to prevent similarity ratings from being affected by differences in consonant production, the stimuli were edited to remove traces of consonant articulation. Thus the portion corresponding to the /b/ closure and the release of the /t/ were excluded while maintaining intact the cues to the vowel. Stimuli were normalized for intensity (70dB) to minimize talker-related loudness differences. The resulting stimuli were combined to create a variety

of pairs (see below) with an interstimulus interval of 1.2 seconds.

2.2. Rated dissimilarity task design

The inclusion of all possible combinations of Catalan vowels and English vowels in the experiment would have resulted in a very large number of pairs and consequently a very long task, prone to fatigue effects. A decision was made to select a subset of vowel combinations that included pairs of acoustically close native and non-native vowels as well as pairs involving more distant vowels. The selection of pairs was also guided by the need to construct a balanced design in which all individual vowels appeared the same amount of times throughout the task, in both experiments. In order to achieve this balance, every Catalan and every English vowel appeared in the same number of combinations across the total number of same-language and different-language pairs. This resulted in 18 Catalan-English pairs that were used in both experiments and a set of same-language (L1-L1) pairs that were specific to each experiment. Table 1 presents the list of the crucial 18 Cat.-Eng. pairs used in both experiments.

The set of same-language pairs included same-category pairs (e.g., Cat /e/-/e/, /a/-/a/, /i/-/i/, /ei/-/ei/, /ai/-/ai/; Eng /i/-/i/, /ε/-/ε/, /æ/-/æ/, /ei/-/ei/, /ai/-/ai/) and different-category pairs (e.g., Cat /e/-/i/, /ai/-/a/, /ei/-/e/, /ε/-/e/; Eng /i/-/ɪ/, /ε/-/æ/, /ei/-aɪ/, /ɜ/-/ɪ/). Each stimulus pair was presented in three different talker combinations, and in two possible orders (e.g., Cat. /i/-Eng. /i/ and Eng. /i/-Cat. /i/). The resulting total number of trials was 174.

Table 1: Common Catalan/English vowel pair stimuli.

Cat. V - Eng. V.	Cat. V - Eng. V
/a/ /æ/	/ε/ /æ/
/a/ /ε/	/ε/ /ε/
/a/ /ɜ/	/ε/ /ɜ/
/ai/ /ei/	/e/ /ε/
/ai/ /aɪ/	/e/ /ei/
/ai/ /æ/	/e/ /ɪ/
/ei/ /ei/	/i/ /aɪ/
/ei/ /ɪ/	/i/ /ɪ/
/ei/ /i/	/i/ /i/

2.3. Procedure

In each trial listeners were presented with a pair of stimuli and were instructed to indicate on a 9-point scale (1 = very similar, 9 = very dissimilar) the

degree of similarity between the two stimuli. Participants were instructed to use all the values on the scale. A practice task containing a variety of pairs was performed first to ensure familiarity with the task and expose participants to the kind of variability in the actual test.

The listeners performed the task individually in a soundproof booth at Universitat Autònoma de Barcelona (Catalan participants) and University College London (English participants). They listened to the stimuli over headphones and gave their responses by clicking on a rating scale appearing on a computer screen. Praat software [2] was used to conduct the experiments. Participants took approximately 20 minutes to complete the task.

2.4. Participants

2.4.1. Experiment 1. Native English speakers

Twenty-seven native speakers of Southern British English participated in the study (four males and 23 females, average age of 22 years). Most participants were undergraduate and graduate students. They had no knowledge of Catalan and their knowledge of foreign languages was limited.

2.4.2. Experiment 2. Native Catalan speakers

The Catalan speakers were 27 Catalan-dominant Catalan-Spanish bilinguals (12 males and 15 females) selected from a pool of 38 potential participants based on their personal and linguistic background. Their average age was 24. Most participants were undergraduate students. The ideal participants would be speakers with little or no knowledge of the non-native language in the experiment. While this was not a problem for the English subjects, it was more difficult to accomplish with the Catalan participants. Most had studied English as a foreign language (EFL) at school. However, the focus of EFL instruction is typically grammar and vocabulary rather than pronunciation and perception. None of the Catalan participants were English students or teachers or had spent more than one month in an English-speaking country. Thus their familiarity with the English vowel system was considered minimal.

All participants reported normal hearing and were compensated for their participation.

3. RESULTS AND DISCUSSION

Table 2 presents the average ratings obtained for each pair of Catalan-English vowels by each group of listeners. The results for the same-category vowel pairs for each listener group are provided in Table 3.

Recall that on the 9-point scale, 1 indicated very similar and 9 meant very different.

Table 2: Ratings provided by the Catalan listeners and the English listeners to each Cat.-Eng. vowel pair and average across the two groups.

Dissimilarity rating			
V pair (Cat.-Eng.)	Catalan Listeners	English Listeners	Average
/ɛ/-/ɛ/	2.3	1.8	2.0
/e/-/ɪ/	2.3	1.8	2.1
/a/-/æ/	2.4	1.8	2.1
/i/-/i/	2.8	2.6	2.7
/ei/-/ei/	4.0	2.4	3.2
/ai/-/ai/	4.1	3.0	3.6
/e/-/ɛ/	5.3	4.3	4.8
/i/-/ɪ/	5.6	3.9	4.8
/ai/-/ei/	5.7	4.2	4.9
/ei/-/i/	5.9	4.7	5.3
/ei/-/ɪ/	5.5	5.8	5.7
/ai/-/æ/	5.5	6.3	5.9
/e/-/ei/	6.4	5.6	6.0
/a/-/ɛ/	7.1	6.1	6.6
/ɛ/-/æ/	7.4	6.1	6.8
/a/-/ɜ/	6.3	7.7	7.0
/ɛ/-/ɜ/	6.9	7.4	7.2
/i/-/ai/	8.0	6.9	7.4
Mean	5.2	4.6	

On the whole, the two groups appeared to judge the degree of similarity for Catalan-English vowel pairs in a similar fashion. The six pairs that obtained the lowest dissimilarity ratings were the same for the two groups of listeners (see Table 2). Despite the similarity of the results, the two groups performed differently, as indicated by the significant result of a paired samples t-test ($t = 2.65, p = .008$). This difference is due to the overall lower dissimilarity ratings given by the English listeners. In other words, the English listeners tended to rate native-nonnative vowel pairs as slightly less dissimilar than the Catalan listeners did (mean ratings of 4.7 and 5.1, respectively). This outcome runs counter to the prediction derived from Dispersion Theory that a smaller vowel inventory, characterized by greater variability within each category, may allow for greater tolerance for non-prototypical exemplars and consequently lower dissimilarity ratings.

This difference between the two groups was also observed in the ratings for same-language pairs, as illustrated in Table 3. Also in this case Catalan listeners tended to give higher dissimilarity ratings than English listeners.

Table 3: Ratings provided by the Catalan listeners and the English listeners to the same-category Cat.-Cat. and Eng.-Eng. pairs.

Catalan same-V pairs		English same-V pairs	
/a/-/a/	2.0	/æ/-/æ/	1.2
/ai/-/ai/	1.9	/ai/-/ai/	1.5
/ei/-/ei/	2.5	/ε/-/ε/	1.4
/ε/-/ε/	1.9	/ei/-/ei/	1.6
/i/-/i/	1.5	/i/-/i/	1.9
Mean	2.0	Mean	1.5

Despite the overall higher ratings by the Catalan listeners, the general pattern of results showed very consistent ratings across pairs for the two listening groups. This pattern was reflected in a test of correlation ($r = .899$). The finding that the same six pairs obtained the lowest dissimilarity ratings by both groups of listeners (Catalan-English /ε/-/ε/, /e/-/ɪ/, /a/-/æ/, /ei/-/ei/, /ai/-/ai/, and /i/-/i/), is consistent with previous studies that reported high degrees of perceptual assimilation of English /æ, ε, i, ei, ai/ to Catalan /a, ε, i, ei, ai/ by Catalan listeners in an interlingual identification experiment [3]. Three of these six pairs received ratings that were within the range of those obtained by same-category L1-L1 pairs (Cat.-Eng. /ε/-/ε/, /e/-/ɪ/ and /a/-/æ/).

This pattern of results suggests that English learners of Catalan or Catalan learners of English may establish equivalence relationships between the closest L1 and TL vowels. According to current L2 models [1,4], these vowels will not be accurately categorized as the equivalence classification will prevent L2 learners from establishing new categories for L2 sounds. However, vowels that received ratings comparable to same-category L1-L1 pairs may pattern as perceptually identical, and direct substitution of a L1 vowel for a TL vowel may go unnoticed by native TL listeners. Other vowels, like English /ɜ/, appear to have no clear match in Catalan. Such vowels are good candidates for accurate L2 categorization given enough exposure to the target language [4]. Comparing the outcomes of the two experiments, the consistency of the results across the two listener groups points to the suitability of a reciprocal measure of cross-linguistic perceptual similarity as a first step in a study of non-native speech perception.

4. CONCLUSION

The perceptual similarity between a subset of the English and Catalan vowel inventories was assessed

by means of a rated dissimilarity task administered to native speakers of both languages. The results showed a very consistent pattern of non-native to native category mapping for both English and Catalan listeners, supporting a reciprocal or bidirectional approach to testing cross-language similarity. The results did not support the idea that a smaller inventory, with presumably greater within-category variability, would result in lower dissimilarity ratings. A subset of the native-nonnative pairs received similarity ratings that were comparable to same-category L1-L1 pairs, showing that some non-native categories may be near-identical to native categories. It is argued that a complete assessment of cross-linguistic similarity is a necessary first step in any L2 speech learning study.

This research was supported by a research grant from the Spanish Ministry of Economy and Competitiveness (FFI2013-46354-P) and by a grant from the Catalan Government (2014SGR61).

5. REFERENCES

- [1] Best C.T., Tyler M.D. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In Bohn, O-S, Munro, M. J. (eds.), *Language Experience in Second Language Speech Learning*. Amsterdam: John Benjamins. 13–34.
- [2] Boersma, P., Weenink, D. 2014. Praat: doing phonetics by computer. <http://www.praat.org/>
- [3] Cebrian, J., Mora, J.C., Aliaga-García, C. 2011. Assessing crosslinguistic similarity by means of rated discrimination and perceptual assimilation tasks. In Wrembel, M., Kul, M., Dziubalska-Kořaczyk, K. (eds.), *Achievements and Perspectives in the Acquisition of L2 Speech*. Frankfurt: Peter Lang. Vol I. 41-52.
- [4] Flege, J. E. 1995. Second language speech learning: Theory, findings and problems. In Strange, W. (ed.), *Speech Perception and Linguistic Experience: Issues in Crosslanguage Research*. Timonium, MD: York Press. 233–277.
- [5] Flege, J. E., Munro, M. J., Fox, R. A. 1994. Auditory and categorical effects on cross-language vowel perception. *J. Acoust. Soc. Am.*, 95, 3623–3641.
- [6] Lindblom, B. 1986. Phonetic universals in vowel systems. In Ohala, J., Jaeger, J. (eds.), *Experimental Phonology*. Academic New York. 13-44.
- [7] Strange, W. 2007. Cross-language phonetic similarity of vowels. Theoretical and methodological issues. In Bohn, O-S, Munro, M. J. (eds.), *Language Experience in Second Language Speech Learning*. Amsterdam: John Benjamins. 35–55.