Cross-language perceptual assimilation and discrimination of Southern British English vowels by Greek and Japanese learners of English

2 AUTHORS:


Angelos Lengeris
University of Cyprus


Valerie Hazan
University College London

CROSS-LANGUAGE PERCEPTUAL ASSIMILATION AND DISCRIMINATION OF SOUTHERN BRITISH ENGLISH VOWELS BY GREEK AND JAPANESE LEARNERS OF ENGLISH

Angelos Lengeris and Valerie Hazan

Department of Phonetics and Linguistics, University College London, London, UK
a.leggeris@ucl.ac.uk, v.hazan@phon.ucl.ac.uk

ABSTRACT

This study compared the perception of Southern British English (SBE) vowels by Greek (Gr) and Japanese (J) learners of English, first using a cross-language mapping task to make predictions regarding SBE vowels' discriminability and then a categorial oddity discrimination task to test these predictions. Both G and J listeners have five vowel qualities in their inventories but only Japanese has short/long versions of each quality. Differences between L2 groups in both experiments support the view that L2 perception cannot accurately be predicted by an abstract phonological cross-language comparison. L1 experience with duration helps J listeners in perceiving some SBE contrasts; Gr listeners nonetheless seem to be able to use duration to some extent despite the lack of experience with the particular feature in L1.

Keywords: L2 vowel perception, cue weighting, duration.

1. INTRODUCTION

Second language (L2) perception is generally a difficult task for learners of a new language. Best's [1] Perceptual Assimilation Model (PAM) posits that non-native categories assimilate into the native ones and that discrimination of a non-native contrast can be predicted quite accurately by the perceived phonetic similarity between the sounds in the contrast and the L1 categories. However, there are a limited number of studies providing both assimilation and discrimination results, e.g. [3], and only one [5] that directly compared two L2 groups when perceiving non-native vowels using the above method. Furthermore, the majority of studies in L2 vowel perception focus on very few contrasts (e.g. the English /iː-ɪ/ distinction) instead of examining complete vowel inventories.

Another important issue concerns the relative weighting of acoustic cues by L2 learners who may differ from the L1 norm. Previous research has shown that when L2 learners are unable to differentiate L2 vowel contrasts based on spectral cues, they rely on temporal ones, irrespective of whether they use duration in their L1, the reason being that duration is a cue that is salient and easy to access [2]. According to the "feature" hypothesis [6] on the other hand, category formation is difficult when based on a phonetic feature/cue not used in L1 to signal phonological contrasts.

By directly comparing the perception of the same vowel system by two L2 groups with (phonologically) same quality distinctions /i e a o u/ that only differ in the use of duration in L1 we can test the degree to which each group uses duration. By using bVb and bVp contexts for the stimuli we can examine the effect of vowel intrinsic duration on both L2 groups' performance.

To assess L2 vowel perception we first used a cross-language mapping task with goodness ratings [1, 3] to make specific predictions in terms of discrimination difficulty for several SBE contrasts varying in degree of difficulty; we then used a categorial discrimination task to evaluate these predictions.

2. METHOD

2.1. Assimilation test

2.1.1. Participants

Seven native speakers of Standard Athenian Modern Greek (4 f, 3 m) aged 18-25 and seven native speakers of Standard Tokyo Japanese (4 f, 3 m) aged 18-23 were tested. The two groups had studied English in Greece and Japan respectively for 10-15 years but had very little conversational experience with native speakers of English and had not spent a period of more than 2 weeks in an English-speaking environment. None of the participants was experienced in any other foreign language apart from English. Seven native speakers of SBE (5 f, 2m) aged 18-28 served as controls.

2.1.2. Stimuli

Three native speakers of SBE (all f) born in London and aged 23-30 uttered the stimuli. Each speaker produced four randomized (blocked by context) repetitions of the eleven SBE monophthongs in a bVb and bVp context in the sentence *I read ____ on the screen*. The best three tokens for each vowel were used (3 speakers * 11 vowels * 2 contexts * 3 repetitions = 198 stimuli in total). The stimuli were edited out of the sentences and normalized for peak amplitude.

All the recordings took place in the UCL anechoic room using a DAT recorder at a 44.1 kHz sampling rate and were then down-sampled to 22.05 kHz.

2.1.3. Procedure

Participants were tested individually in quiet rooms; the Gr listeners in Greece and the J listeners in the UK during their two-week stay there for the UCL Summer Course in Phonetics. They all completed a language background questionnaire.

Listeners were presented the stimuli at a comfortable sound level via high quality headphones and categorized each vowel token in terms of their native vowel categories (by clicking the relevant option on a computer screen). Then they heard the same token again and rated its goodness-to-fit to the just chosen L1 category in a scale from 1 (= totally different) to 7 (= identical).

2.2. Discrimination test

2.2.1. Participants and stimuli

The participants and the stimuli were the same as in experiment 1.

2.2.2. Procedure

A categorial oddity discrimination test [3] was used to test the discrimination of nine SBE contrasts /iː-ɪ/, /ɪ-e/, /æ-ʌ/, /æ-aː/, /æ-ɜː/, /ʌ-aː/, /ɒ-ɔː/, /ʊ-uː/ and /ɔː-uː/, chosen to vary in degree of difficulty. The contrasts were presented in triads with each stimulus spoken by a different speaker. Each contrast was tested by eight 'different' trials that contained an odd vowel category and eight 'catch' trials that contained three tokens of the same vowel category (always physically different).

The participants were instructed to identify the odd item out by clicking "1", "2" or "3" (in the 'different' trials) or "same" (in the "catch" trials). To minimize response bias, *A'* scores [7] were computed for each contrast based on 'hits' (the odd item was correctly selected) and 'false alarms' (an item was incorrectly selected in a "catch" trial). *A'* score of 1.0 indicates perfect discrimination whereas *A'* score of 0.5 indicates discrimination at chance level.

3. RESULTS

3.1. Assimilation test

Table 1 presents the perceived relationship between SBE vowels and Gr (shaded cells) and J (non-shaded cells) vowels. Due to space limitations, only the results for bVb context are given in detail. The results for bVp context are given (in parentheses) only when there are notable differences across contexts (<5% difference). Consistency in the most frequent response ranged from 56%-100% (ratings from 3-5) for Gr listeners and 47%-96% (ratings from 3-6) for J listeners. Spectral assimilation patterns were similar across L2 groups for most SBE vowels. However, there were some differences between Gr and J listeners' responses (shown in boldface) despite the spectral similarity of their vowel systems.

Focusing on the most frequent response and disregarding duration, SBE /ɜː/ assimilated into Gr /e/ and J /a/ respectively and SBE /aː/ assimilated into Gr /o/ and J /a/ respectively. According to PAM [1], these differences lead to different predictions for Gr and J listeners' L2 discrimination performance when these two vowels are paired with others to constitute L2 contrasts.

In general, Gr listeners were expected to obtain: excellent scores for /ɪ-e/ and /æ-ɜː/, both *two-category (TC)* contrasts according to PAM terminology; very good scores for /æ-aː/, /ʌ-aː/, /ɒ-ɔː/ and /ɔː-uː/, all *Uncategorizable* vs. *Categorized* (*UC)* contrasts, where one vowel of the contrast falls between two L1 categories; very good scores for /æ-ʌ/, a *Category Goodness (CG)* contrast; and low scores for /iː-ɪ/ and /ʊ-uː/, both *Single Category (SC)* contrasts. Context does not seem to affect Gr listeners' assimilation patterns.

J listeners were expected to obtain excellent scores for several SBE contrasts /iː-ɪ/, /ɪ-e/, /ʊ-uː/, /ɔː-uː/ and /ɒ-ɔː/, all *TC* contrasts across contexts. However, assimilation patterns and, consequently, PAM predictions for the other SBE contrasts, /æ-ʌ/, /æ-aː/, /æ-ɜː/ and /ʌ-aː/, vary as a function of

Table 1: Most frequent and second most frequent response, percentage of total opportunities and median goodness ratings by Gr and J listeners for SBE vowels in bVb and bVp context (see text for details).

		1st response			2nd response		
L2	L1		%	Rating		%	Rating
iː	Gr	i	100	5			
	J	ii	82	5 (6)	i	12	5
ɪ	Gr	i	100	5			
	J	i	72	4	e	26	4
e	Gr	e	97	5			
	J	e	93	5	ee	6	4
ɜː	Gr	**e**	**77 (67)**	**3**	**o**	**8 (25)**	**2**
	J	**aa**	**50 (47)**	**3**	**a**	**19 (28)**	**2**
æ	Gr	a	95	5	e	5	1
	J	a	64 (84)	3 (4)	aa	35 (16)	3
ʌ	Gr	a	65	4	o	35	5
	J	a	89	4	o	6	3
ɒ	Gr	o	97	5			
	J	o	89	5 (4)	oo	6	3
aː	Gr	**o**	**57**	**4**	**a**	**43**	**3**
	J	**aa**	**64 (53)**	**5 (4)**	**a**	**25 (20)**	**4 (3)**
ɔː	Gr	o	56	4	u	44	3
	J	oo	69 (74)	5	o	21	5
ʊ	Gr	u	92	4	o	6	2
	J	u	85	4	uu	10	5
uː	Gr	u	86 (92)	4 (3)	i	14	2
	J	uu	85	5	u	10	5

context. In bVb context, SBE /æ/ assimilated into J /a/ 64% of instances (median rating 3) and J /aa/ 35% of instances whereas in bVp context 84% (median rating 4) and 16% respectively. Therefore, /æ-ʌ/ is a *CG* contrast in the former context and a *SG* contrast in the latter context (SBE /ʌ/ assimilated into J /a/ category with a median rating of 4 across contexts). Inversely, vowels in /æ-aː/ and /æ-ɜː/ were more similar before a voiced contrast and therefore more difficult to discriminate (it is difficult to categorically assign these contrasts into PAM categories but they certainly seem very similar for J listeners with overlapping between 1st response for one vowel and 2nd response for the other in each contrast). Finally, discrimination of /ʌ-aː/ is expected to be easier for J listeners in bVb than in bVp context.

3.2. Discrimination test

Fig. 1 and 2 present the range of A' scores obtained by Gr and J listeners in bVb and bVp context respectively. Native speakers of English obtained excellent discrimination scores (A' <0.9) for all contrasts so their results are not presented graphically but are nevertheless included in the statistical analysis. A' scores were submitted to a

Figure 1: *A'* scores by SBE contrast in bVb context.

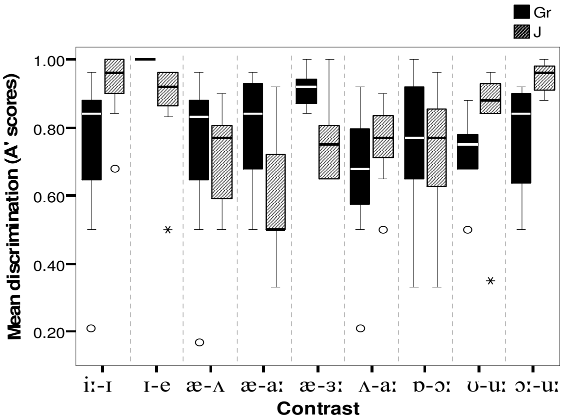
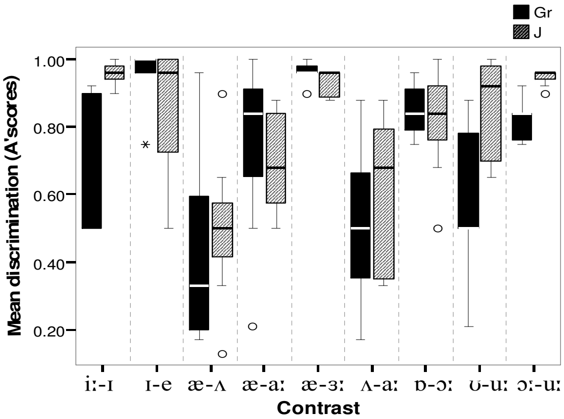


Figure 2: *A'* scores by SBE contrast in bVp context.



mixed ANOVA with (3) group as a between-subjects factor and (9) contrast and (2) context as within-subjects factors. Two of the three main effects reached significance [group $F(2,18) = 20.087$, $p<0.001$; contrast $F(8,114) = 9.303$, $p<0.001$] as did all three two-way interactions [contrast × group $F(16,144) = 3.115$, $p<0.001$; contrast × context $F(2,18) = 20.087$, $p<0.001$; context × group $F(2,18) = 7.259$, $p<0.01$].

The simple effect of group was significant for all SBE contrasts at a significant level of $p<0.01$. A Tukey's test revealed that NE listeners received higher scores than Gr listeners for all contrasts but /ɪ-e/ and /æ-ɜː/ (*TC* contrasts) and higher scores than J listeners for /æ-ʌ/, /æ-aː/, /æ-ɜː/, /ʌ-aː/, /ɒ-ɔː/ but not for /iː-ɪ/, /ɪ-e/, /ʊ-uː/, /ɔː-uː/ (*TC* contrasts). Consequently, within L2 groups, Gr listeners received higher scores than J did for /æ-ɜː/ and lower for /iː-ɪ/, /ʊ-uː/ and /ɔː-uː/.

Inspection of Fig. 1 and 2 reveals that the contrast × context interaction occurred because of the effect of context not being uniform across contrasts. The context × group interaction occurred because NE and J listeners had similar overall scores across contexts (note though that the latter group had, as predicted, significantly higher scores for /æ-ʌ/ and /ʌ-aː/ in bVb context and inversely higher scores for /æ-aː/ and /æ-ɜː/ in bVp context) whereas Gr listeners had lower overall scores in bVp context.

4. DISCUSSION

This study examined the extent to which L1 spectral and durational experience affects Gr and J listeners' perceptual assimilation and discrimination of SBE vowels. Experiment 1 revealed some cross-language differences in perceptual assimilation patterns which could not be predicted by an abstract phonological comparison of the Gr and J vowel systems.

Results in Experiment 2 confirmed most of PAM predictions. Both L2 groups obtained native-like *A'* scores for *TC* contrasts and lower scores for *UC*, *CG* and *SC* contrasts. In addition, Gr listeners were better at /æ-ɜː/ whereas J listeners were better at /iː-ɪ/, /ʊ-uː/ and /ɔː-uː/ again as predicted by their assimilation results. In two cases though, Gr listeners' discrimination scores were not consistent with their assimilation results. Importantly, both instances are related with the main question asked by the present study, namely whether and to what extent Greek speakers have access to duration information when perceiving the English vowels.

Firstly, Gr listeners obtained relatively high *A'* scores (0.6-0.8) for /iː-ɪ/ and /ʊ-uː/, both *SC* contrasts. Similar results for *SC* contrasts are reported in [4]. It is reasonable to assume that Gr listeners used duration to some extent to discriminate these contrasts although not as effectively as J listeners who clearly assigned the English tense and lax vowels to their long and short L1 categories respectively. However, Gr listeners did not show similar sensitivity to duration in experiment 1 (they gave practically the same goodness ratings to both tense and lax vowels).

Secondly, Gr listeners obtained lower *A'* scores in bVp context, which is further evidence that they are sensitive to duration information, but again that was not shown in Experiment 1 (they gave similar ratings across contexts). We believe that for both findings the reason is the same: Gr listeners paid more attention to spectral cues in Experiment 1 because of the nature and/or format of the task; they were asked to judge vowel similarity using a single scale. For a naïve listener with no duration experience it is very unlikely to base her judgment on temporal instead of spectral similarity even if still able to perform quantity distinctions. In fact, during testing two participants did ask whether they should judge similarity with respect to quality or quantity. In contrast, the discrimination test used in Experiment 2 must have prompted Gr listeners to respond to duration differences between contiguous tokens especially in cases where they were unable to detect spectral ones.

Overall, listeners with no L1 durational experience have access to duration information in L2. In that respect, Bohn's hypothesis [2] is supported. Still, Gr listeners obtained lower *A'* scores than J listeners. In that respect, the "feature" hypothesis [6] cannot be readily rejected. To test L2 groups with no L1 duration experience we may have to modify the assimilation experiment paradigm to allow them to demonstrate their sensitivity to temporal cues.

5. ACKNOWLEDGMENTS

We would like to thank Michael Ashby for his help in stimuli selection.

6. REFERENCES

[1] Best, C. T. 1995. A direct realist view of cross-language speech perception. In W. Strange (eds), *Speech perception and linguistic experience: Issues in cross-language research*. Timonium, MD: York Press, 171–204.

[2] Bohn, O.-S., 1995. Cross-language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (eds), *Speech perception and linguistic experience: Issues in cross-language research.* Timonium, MD: York Press, 279–304.

[3] 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, 2711-2723.

[4] Harnsberger, J. D. 2000. On the relationship between identification and discrimination of non-native nasal consonants. *J. Acoust. Soc. Am.* 110, 489-503.

[5] Ingram, J., Park, S-G. 1997. Cross-language vowel perception and production by Japanese and Korean learners of English *Journal of Phonetics* 25, 343-370.

[6] McAllister, R., Flege, J. E., Piske, T. 2002. The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics* 30, 229–258.

[7] Snodgrass, J., Levy-Berger, G., Haydon, M. 1985. *Human experimental psychology*. Oxford: Oxford University Press.