



The relation between perception and production in L2 phonological processing

Sharon Peperkamp¹, Camillia Bouchon^{1,2}

¹Laboratoire de Sciences Cognitives et Psycholinguistique, Paris, France

²Laboratoire Psychologie de la Perception, Paris, France

Sharon.Peperkamp@ens.fr, camillia.bouchon@etu.parisdescartes.fr

Abstract

Seventeen French-English bilinguals read aloud a set of English sentences and performed an ABX discrimination task that assessed their perception of the English /ɪ/-i/ contrast. Global nativelikeness in production correlated with pronunciation accuracy for the vowels /ɪ/ and /i/, and both production measures correlated with self-estimated pronunciation skills. However, performance on the perception task did not correlate with either global nativelikeness or /ɪ,i/ pronunciation accuracy. These results are discussed in light of theories about the relation between perception and production in L2 phonological processing.

Index Terms: speech perception, speech production, bilingualism, L2, phonology, phonetics, English vowels

1. Introduction

The relation between perception and production in non-native phonological processing has been investigated with a variety of methods in various types of population. Some of the debate has centered around the question of whether there is a causal relationship between performance in one modality and that in the other. According to the Speech Learning Model, the accuracy with which non-native sounds are produced is limited by how accurately they are perceived, and an increase in performance in production is necessarily preceded by one in perception [1]. Experiments with L2 learners have yielded evidence both for and against this hypothesis [2,3]. A problem with this research question is that it is particularly difficult to compare performance in a perception experiment with that in a production experiment, due to differences in the experimental methods used for the two modalities.

An alternative way to investigate the relation between perception and production is to focus on the question of whether perception and production are correlated. Several studies with bilinguals have found a moderate correlation between the perception and the production of L2 segmental contrasts [4,5,6/7,8,9,10]. They are paralleled by data from native language processing that likewise show a correlation between perception and production in accent change in young adults [11]. However, the correlation seems to be restricted to certain measures [12], and perception and production skills for L2 contrasts develop differently [13/14]. Moreover, measures of white matter anatomy in language-related brain areas show a partial dissociation between perception and production of novel contrasts [15/16]. These results are in accordance with research with brain-lesioned patients, who can be selectively impaired in either perception [17] or production [18].

In this paper, we investigate the relation between L2 perception and production in French-English bilinguals who have learned English in school and who live in France but regularly use English. Participants perform both a sentence reading task and an ABX discrimination task. The production

task yields a global nativelikeness score as well as an accuracy score for the pronunciation of the vowels /ɪ/ and /i/; the /ɪ/-i/ contrast is particularly difficult for native speakers of French, because their language does not have /ɪ/ (or, for that matter, any other lax vowel). The perception task assesses the discrimination of the same tense-lax vowel contrast. The relation between production and perception is examined by means of correlation analyses.

2. Sample

In order to select French-English bilinguals who speak English with varying degrees of a French accent, we used two recruitment adds, one that did and one that did not mention a requirement to speak English without a noticeable foreign accent. We also approached two bilinguals whom we regarded to have an excellent English pronunciation. Overall, we selected seventeen bilinguals in the Paris region, 11 men and 6 women aged between 22 and 34 (mean: 26). They had all started to learn English in school between the ages of 7 and 13 (mean: 11 years). They had all been exposed mainly to British English and used English on a daily basis.

The bilinguals filled in a questionnaire in which they evaluated on a scale from 1 to 10 their English pronunciation (mean: 6.5), grammar (mean: 7.5), vocabulary (mean: 7.1) and orthography (mean: 7.8), and indicated how much importance they attached to having a good English pronunciation (mean: 6.7).

3. Experiment

3.1. Method

3.1.1. Participants, stimuli and procedure for the production task

All bilinguals, as well as four English speakers who served as controls, were recorded while reading eight English sentences (see Appendix) in a sound-proof booth. Among the English speakers, two were British (one man, one woman) and two were Irish (both men). Next, the recordings were scored in two ways:

First, ten native speakers of English (three American and seven British) individually judged the nativelikeness of each of the speakers' accent on a scale from 1 ("very strong foreign accent; definitely nonnative") to 5 ("no foreign accent at all; definitely native"). They listened to all eight sentences of a given speaker and then provided their rating before moving on to the next speaker. The task started with a instruction phase in which the recordings of an additional French-English bilingual with a strong French accent was played, followed by the recordings of two additional native speakers of English with a regional accent, i.e. one Irish speaker, and one British speaker if the judge was American and one American speaker if the

judge was British. It was explained that the French speaker had a strong accent and hence should receive a low rating, and that the native speakers should receive a high rating despite a possible regional accent (since their accent was not foreign). The recordings of the bilingual participants and the four control native English speakers were then presented in a random order.

Second, all tokens of /i/ and /ɪ/ occurring in a stressed syllable of a content word were excised out of the audio files of the 17 bilinguals and the four native English speakers. The sentences contained five such tokens of /i/ and seven of /ɪ/ (see Appendix). Another native speaker of English performed a forced choice identification task on the total of 252 ((17+4)*(5+7)) tokens, presented in a random order. There were three response categories: “vowel of ship”, “vowel of sheep”, and “neither”.

3.1.2. Participants, stimuli and procedure for the perception task

All bilinguals, as well as eight native speakers of English participated. The English speakers, five men and three women aged between 19 and 22 (mean: 21), did not participate in the production task.

Twenty triplets of trisyllabic English pseudo-words ending in *-ing*, *-ish*, *-ous*, *-y*, and *-er* and with stress on the first syllable were selected; members of each triplet differed only in the vowel of the first syllable, which was /i/, /ɪ/, or /u/ (see Appendix). The items were recorded by three native speakers of British English, two men and one woman.

Participants performed an ABX discrimination task in which experimental trials concerned the contrast /ɪ/-/i/ and control trials the contrast /i/-/u/. Each trial consisted of the presentation of three stimuli belonging to the same triplet (A, B and X), with an ISI of 500ms. A and B were always produced by the two male speakers and X by the female speaker. Participants had to press a button on their left or right to indicate whether X was the same as A or as B, respectively. The trial ended immediately after the response was given or after 3500ms had elapsed, whichever came first. The next trial started 1000ms later.

The experiment started with a warm-up phase of 20 trials, during which participants received feedback as to whether their responses were correct. In the case of an incorrect response or no response within 3500ms the trial was repeated until the correct response was given.

The test phase consisted of 160 trials, eight per triplet, divided over two blocks. In each block, half of the trials concerned the experimental contrast, the other half the control contrast; the identity of X as well as the correct response (A or B) were counterbalanced and the trials were presented in a random order.

3.2. Results and discussion

3.2.1. Production

Tokens of /ɪ/ nor /i/ that were identified by the native English listener as neither one of these vowels were considered as incorrect pronunciations. This concerned 4.5% of the tokens produced by the bilinguals (/ɪ/: 6.0%, /i/: 3.3%) and none of those produced by the native English participants.

The mean nativelikeness scores and the mean percentages of accurate pronunciation of /ɪ/ and /i/ for the bilingual and the English participants are shown in Table 1.

Table 1. Results of production task. Standard errors are shown in parentheses.

	bilinguals	English
nativelikeness (on a scale from 1 to 5)	2.7 (0.2)	4.9 (0.1)
accuracy (%)		
/i/	51.4 (6.5)	100
/ɪ/	83.5 (4.9)	95.0 (5.0)
mean	67.5 (4.9)	97.5 (2.5)

The nativelikeness scores were submitted to an ANOVA with the factors Group (bilinguals vs. English) and Judge (American vs. British), which yielded an effect of Group only ($F(1,19)=14.0$, $p<.001$). The absence of an effect of Judge shows that the participants’ accents were not rated differently by the American and the British judges.

Concerning the accuracy data, note that the English participants showed no variance in their performance on /ɪ/. Accuracy was therefore analyzed in two restricted one-way ANOVAs. The first of these showed that the bilinguals performed better on /i/ than on /ɪ/ ($F(1,16)=20.5$, $p<.0001$). The second one showed that the English and the bilingual participants performed equally well on the vowel /i/ ($p>.1$).

Figure 1 shows a scatterplot of the individual production scores on both measures. Note that despite the fact that about half of the bilingual participants had been recruited by means of an add that required native-like pronunciation, the scores of few of them are close to those of the English controls.¹ The two production scores were highly correlated (bilinguals only: $r^2=.65$, $p<.0001$; all participants: $r^2=.78$ $p<.0001$).

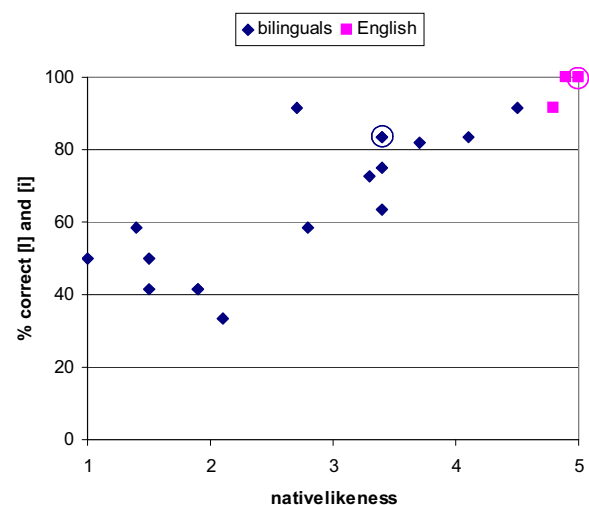


Figure 1: Scatterplot with individual scores for two production measures. The two encircled datapoints represent the data of two bilinguals and of two English participants, respectively.

Finally, we carried out a series of linear regressions on the individual bilinguals’ production scores with their scores in the questionnaire as regressors.² As can be seen in Table 2,

¹ It is true, though, that all but one of the participants who had replied to the neutral add scored lower on both measures than those who had replied to the add that focused on native-like pronunciation.

² Due to the relative homogeneity in this variable, age of acquisition was not used as a regressor.

nativelikeness correlates significantly with four out of the five regressors, /i,i/ pronunciation accuracy with three.

Table 2. Results of linear regressions.

	nativelikeness		% correct /i,i/	
	r ²	p	r ²	p
self-estimated pronunciation	.69	.0001	.44	.004
self-estimated grammar	.49	.002	.40	.007
self-estimated vocabulary	.22	.06	.12	n.s.
self-estimated orthography	.43	.004	.36	.02
motivation for pronunciation	.50	.002	.44	.004

Separate multiple stepwise regression analyses on the two production scores rejected all regressors except self-estimated pronunciation. Hence, once self-estimated pronunciation is taken into account the other regressors do not add any more predictive power for either the nativelikeness score or the /i,i/ pronunciation accuracy rate.

Taken together, these results show - not surprisingly - that the bilingual participants spoke English with a foreign accent, and that their pronunciation of /i/, which is not used in French, was less accurate than that of /i/. The correlation that was observed between these two measures is in agreement with results reported in previous research [19, and references cited therein]. There was substantial inter-individual variation, and the bilinguals' individual scores were best predicted by their self-estimated pronunciation scores.

3.2.2. Perception

The mean percentages of errors by group and by contrast for the perception task are shown in Table 3.

Table 3. Mean percentages of errors in perception task. Standard errors are shown in parentheses.

	/i/-/i/		/i/-/u/	
	Mean	SE	Mean	SE
bilinguals	31.7%	(2.4)	18.7%	(1.9)
English	11.2%	(1.8)	12.0%	(1.4)

The error rates were submitted to two ANOVAs, one by participants and the other by items, with the factors Group (bilinguals vs. English) and Contrast (/i/-/i/ vs. /i/-/u/). These ANOVAs revealed main effects of Group ($F_1(1,23)=20.1$, $p<.0001$; $F_2(1,19)=69.0$, $p<.0001$) and Contrast ($F_1(1,23)=38.8$, $p<.0001$; $F_2(1,19)=7.6$, $p<.02$), as well as an interaction between these two factors ($F_1(1,23)=21.7$, $p<.0001$; $F_2(1,19)=59.4$, $p<.0001$). The interaction was due to the fact that the bilinguals made more errors on the /i/-/i/ than on the /i/-/u/ contrast ($F_1(1,16)=44.8$, $p<.0001$; $F_2(1,19)=31.4$, $p<.0001$), whereas the English showed no difference ($F_{1,2}<1$).

These results show that French-English bilinguals, as opposed to native speakers of English, have difficulties perceiving the contrast between /i/ and /i/.³

³ Analyses restricted to the first and the second block, not shown here for reasons of space, yielded the same patterns of results, attesting to the reliability of the overall results. Analyses of the reaction times also yielded very similar results. In particular, there was an interaction between Group and Contrast, due to slower performance of only the bilinguals on the /i/-/i/ contrast.

3.2.3. Correlations between perception and production

Figure 2 shows the individual scores of the bilingual participants on the two contrasts in the perception task, in order of worst to best performance on the control contrast, i.e. /i/-/u/.

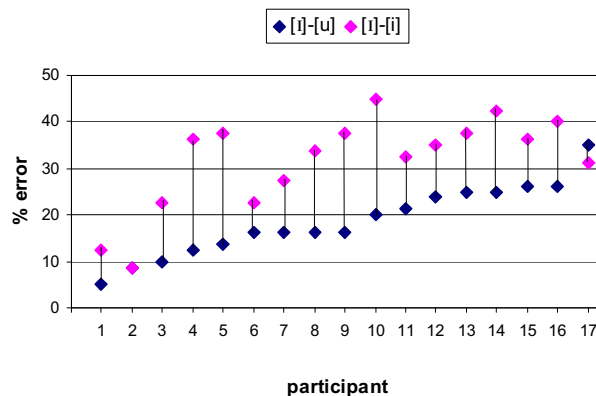


Figure 2: Individual scores of the bilingual participants in the perception task.

Note that the large majority of bilinguals made more errors on the /i/-/i/ contrast than on the /i/-/u/ contrast; the only exceptions are one participant (n° 2) who had the same, low, error rate on both contrasts, and another one (n° 17) who had the highest error rate on the /i/-/i/ contrast.⁴

In order to investigate whether perception and production are correlated in our bilinguals participants, we carried out linear regressions between the two production scores on the one hand and the error rate on the /i/-/i/ contrast in the perception task on the other hand. In both regressions, the error rates on the /i/-/u/ contrast in the perception task were entered as a covariate. These analyses showed that perception and production are not correlated (nativelikeness vs. /i/-/u/ perception: $r^2=-.06$, $p>.1$; /i,i/ production vs. /i/-/i/ perception: $r^2=0$).⁵

4. Conclusions

Testing French-English bilinguals in both a production and a perception task, we found no signs of a correlation between the perception and the production of the English /i/-/i/ contrast (nor between /i/-/i/ discrimination and global nativelikeness in production). This finding contrasts with those of previous studies with L2 learners - several of which likewise examining vowel contrasts - that observed a moderate correlation between perception and production [4,5,6/7,8,9,10]. What might account for this discrepancy?

Clearly, the lack of a correlation in the present results is not due to a lack of variance in either the production or the perception scores; indeed, we obtained large ranges of scores

⁴ The native English showed a rather different pattern: three made more errors on the /i/-/i/ contrast, one had equal error rates, and four made more errors on the /i/-/u/ contrast.

⁵ Similarly, no significant correlations were obtained in linear regressions that used individual difference scores for perception, defined as the error rate on the /i/-/i/ contrast minus that on the /i/-/u/ contrast (all $F<1$).

on all measures. Moreover, the strong correlation between the two production scores, as well as the correlations between each of the production scores and self-estimated pronunciation, attest of the reliability of the production measures. Likewise, the block analyses for the discrimination task (see footnote 3) shows the reliability of the perception measure.

We hypothesize that the lack of a correlation is rather due to the fact that for perception we used a speeded ABX discrimination paradigm with a small ISI of 500ms. In this task, participants cannot loop through their production module when processing the stimuli and making a decision as to their response. By contrast, all the previous studies that found a perception-production correlation used either an off-line task, such as identification or goodness rating, or discrimination with a much larger ISI (at least 1.2s). Thus, we suggest that the correlation between perception and production in L2 phonological processing is dependent upon the possibility to use the phonological loop during perception, and in particular, to subvocally rehearse the stimuli [20]. This hypothesis can be tested in further research that tests the role of the phonological loop in perception experiments, for instance by using articulatory suppression or by manipulating the ISI.

5. Appendix

Materials for the **production** task (sentences 1-6 are from [21], 7 and 8 were newly constructed); the vowels that were excised for the /i/-/ɪ/ judgments are shown in red:

1. Arthur will **finish** his **thesis** within **three weeks**.
2. My **sister** Paula prefers coffee to **tea**.
3. The lad was mad about his dad's new **fad**.
4. Mat's flat is absolutely **fantastic**.
5. It's a **pity** we didn't go to the **city**.
6. You'd better look it up in a **cookbook**.
7. In Johannesburg we **visited** a **township**.
8. You should be good with this cute **little sheep**.

Materials for the **perception** task:

/i/	/ɪ/	/u/
dreederling	driddering	droodering
dreefalous	driffelous	droofelous
dreekerling	drickerling	drookerling
dreeleerous	drillerous	droolerous
dreeterous	dritarous	drooterous
pleeferish	pliferish	plooferish
pleekelish	plickelish	plookelish
pleelerish	plillerish	ploolerish
preetering	pritering	prooterling
sfeeterer	sfitterer	sfooterer
smeederer	smidderer	smooderer
smeefering	smiffering	smoofering
smeelering	smillering	smoolering
spreedely	spriddely	sproodely
spreetely	sprittely	sprootely
sweefery	swiffery	swoofery
sweekely	swickely	swookely
sweelery	swillery	swoolery
treederish	tridderish	trooderish
vleekerous	vlickerous	vlookerous

6. Acknowledgements

Research for this article was supported by the Agence Nationale de la Recherche. We would like to thank Bria Long

for help with the participant recruitment and Charlotte Jacquemot for comments and discussion.

7. References

- [1] Flege, J., "Second-language speech learning: Theory, findings, and problems.", in W. Strange [Ed], *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 233-277, York Press, 1995.
- [2] Llisterri, J., "Relationships between Speech Production and Speech Perception in a Second Language", *Proc. of the 13th Int. Congress of Phonetic Sciences*, Vol. 4, 92-99, 1995.
- [3] Flege, J., "The Relation Between L2 Production and Perception.", *Proc. of the 14th Int. Congress of Phonetic Sciences*, Vol. 2, 1273-1276, 1999.
- [4] Flege, J., "Production and Perception of a Novel, Second-language Phonetic Contrast", *JASA*, 93(3): 1589-1608, 1993.
- [5] Flege, J., MacKay, I., and Meador, D., "Native Italian Speakers' Perception and Production of English Vowels", *JASA*, 106(5):2973-2987, 1999.
- [6] Schmidt, A. and Flege, J., "Effects of Speaking Rate Changes on Native and Non-native Production", *Phonetica*, 52(1):41-54, 1995
- [7] Flege, J., and Schmidt, A., "Native Speakers of Spanish Show Rate-dependent Processing of English Stop Consonants", *Phonetica*, 52(2): 90-111, 1995.
- [8] Flege, J., Bohn, O.-S., and Jang, S., "Effects of Experience on Non-native Speakers' Production and Perception of English Vowels", *Journal of Phonetics*, 25(4):437-470, 1997.
- [9] Bettoni-Techio, M., Rauber, A., and Koerich, R., "Perception and Production of Word-Final Alveolar Stops by Brazilian Portuguese Learners of English", *Proc. of Interspeech 8:2293-2296*, 2007.
- [10] Kluge, D., Rauber, A., Reis, M., and Hoffmann Bion, R., "The Relationship between the Perception and Production of English Nasal Coda by Brazilian Learners of English", *Proc. of Interspeech 8:2297-2300*, 2007.
- [11] Evans, B. and Iverson, P., "Plasticity in Vowel Perception and Production: A Study of Accent Change in Young Adults", *JASA*, 121(6):3814-3826, 2007.
- [12] Hattori, K. and Iverson, P., "English /r/-/l/ Category Assimilation by Japanese Adults: Individual Differences and the Link to Identification Accuracy", *JASA*, 125(1):469-479, 2009.
- [13] de Jong, K., Silbert, N., and Park, H., "Segmental Generalization in Second Language Segment Identification", *Language Learning*, 59(1):1-31, 2009.
- [14] de Jong, K., Hao, Y.-C., and Park, H., "Evidence for Featural Units in the Acquisition of Speech Production Skills: Linguistic Structure in Foreign Accent", *JoPh*, 37(4):357-373, 2009.
- [15] Golestani, N., Molko, N., Dehaene, S., Le Bihan, D., and Pallier, C., "Brain Structure Predicts the Learning of Foreign Speech Sounds", *Cerebral Cortex*, 17(3):575-582, 2007.
- [16] Golestani, N. and Pallier, C., "Anatomical Correlates of Foreign Speech Sound Production", *Cerebral Cortex*, 17(4):929-934, 2007.
- [17] Praamstra P., Hagoort, P., Maassen, B., and Crul, T., "Word Deafness and Auditory Cortical Function. A Case History and Hypothesis", *Brain*, 114(3):1197-1225, 1991.
- [18] Dronkers, N., "A New Brain Region for Coordinating Speech Articulation", *Nature*, 384:159-161, 1996.
- [19] Flege, J., "The role of subject and phonetic variables", in K. Gruber, D. Higgins, K. Olsen, and T. Wysocki [Eds], *Papers From the 34th Annual Meeting of the Chicago Linguistic Society*, Vol. II, 213-232, Chicago Linguistic Society, 1998.
- [20] Jacquemot, C. and Scott, S., "What is the relationship between phonological short-term memory and speech processing?", *TICS*, 10(11): 480-486, 2006.
- [21] Bongaerts, T., Van Summeren, C., Planken, B., and Schils, E., "Ultimate Attainment in the Pronunciation of a Foreign Language", *SSLA*, 19(4):447-465, 1997.