

# DOES PERCEPTION PRECEDE PRODUCTION IN THE INITIAL STAGE OF FRENCH NASAL VOWEL QUALITY ACQUISITION BY JAPANESE LEARNERS? A CORPUS-BASED DISCRIMINATION EXPERIMENT

Sylvain Detey, Isabelle Racine

Waseda University (SILS), Geneva University (ELCF)  
detey@waseda.jp, Isabelle.Racine@unige.ch

## ABSTRACT

This article presents the results of an AXB discrimination task testing the ability of Japanese beginner learners to distinguish three French nasal vowels /ã ã ã/ and compares them with the results of a corpus-based analysis of their production of the same vowels in two tasks (wordlist repetition and wordlist reading). This perception-production study is embedded in an ongoing 2-year longitudinal study in the framework of an international multi-task corpus-based research programme about the phonetic-phonological systems of non-native varieties of French. Our main results indicate first that the /ã/-/ã/ contrast is better perceived than the /ã/-/ã/ contrast and second that /ã/ obtains a lower discrimination rate when it is contrasted with /ã/, whereas no difference is found between /ã/ and /ã/. These findings are discussed in light of the production results and support the principle according to which perception precedes production in the acquisition of a new phonological contrast.

**Keywords:** perception, production, interphonology, corpus, Japanese learners, French nasal vowels

## 1. INTRODUCTION<sup>1</sup>

Following the seminal work by Polivanov and Trubetzkoi, most researchers in the field of second language (hereafter L2) phonology agree that overall, in the process of learning an L2 system, perception usually precedes production, i.e. learners will more easily accurately identify and discriminate L2 phones than accurately produce these L2 phones. The underlying stance is that, apart from coincidental performances, learners cannot produce accurately what they cannot perceive accurately. This follows the metaphor of the phonological sift, and is reflected in L2 speech perception or learning models such as the *Speech Learning Model* (SLM, [10, 11] or the *Second Language Linguistic Perception Model* (L2LP, [9]). Even though some studies actually reported cases where the opposite pattern was found, most researchers and language

teachers alike tend to assume that the general perception-production order prevails. From a methodological viewpoint though, the literature on L2 phonetic-phonological learning can be divided into at least two sets. The first one is largely rooted in psycholinguistics and usually rely on the results of speech perception or production experiments using strict protocols and signal manipulation, sometimes of very fine phonetic properties. The second one is rather linked to general phonetics and phonology and primarily focuses on L2 speech modelization using data from various production tasks. Over the past 15 years, a few researchers have been trying to bridge those two approaches, especially as the study of an L2 phonetic-phonological system, in order to be cognitively plausible, must ideally examine both sides of the learning process (perception and production) over a certain span of time. Besides, the recent development of corpus phonology [8] should also be integrated in the area of L2 acquisition, and not rely on a qualitatively (one task) and quantitatively (one subject) limited set of data. Moreover, many studies in the field of L2 phonology have been carried out with English as the sole target language, which also limits the scope of the structures under scrutiny [12].

In this paper we report the results of a perceptual study which is coupled with a corpus-based production protocol [5, 6, 19] examining similar structures (nasal vowels) with French as a target language and a population of Japanese beginner learners of French. The perception of French nasal vowels by Japanese learners is an original and potentially instructive topic for at least four reasons: 1) nasal vowels are rather marked phonological structures rarely studied in L2 phonology; 2) the Japanese phonemic inventory does not contain any phonological nasal vowels but includes an underspecified nasal segment symbolized as /N/ with various allophonic realizations and is characterized by its intrinsically moraic prosodic structure; 3) the loanword lexicon in Japanese is rich with English (but also French) borrowings and well-known adaptation rules with a dedicated script in the Japanese writing system (katakana); 4) unlike English (which is taught in secondary and primary

schools in Japan), French education usually starts at university in Japan, offering a window into early learning stages.

In our perception-production study, participants had to perform first the perception experiment and subsequently the production tasks to avoid recognition effects. The results of the production part have already been analyzed and presented [4], and they can be summarized as follows: when asked to produce words including the three nasal vowels /*ã* *õ* *ẽ*/ in a repetition and a reading task, the Japanese learners of French under study managed to pronounce these words with a rather high level of target-like achievement (/*ã*/ = 76.03%, /*õ*/ = 76.14% and /*ẽ*/ = 58.78%), and a perceptual three-dimensional analysis of their productions (nasality, vowel quality and consonantal excrescence) in terms of targetlikeness resulted in the following characterization: i) higher rates for nasality than for quality achievement; ii) little difference between /*ã*/ and /*õ*/ but lower rates for /*ẽ*/ in the reading task, whereas the distinction between the three rates was overall not significant in the repetition task; iii) overall, better productions for nasality, quality and excrescences in the repetition rather than in the reading task.

In the present article, we examine the results of the perception task, in which two contrasts were tested: /*ã*/-/*õ*/ on the one hand and /*ã*/-/*ẽ*/ on the other hand, since most confusions for Japanese learners usually happen between these two pairs. We then compare these results with the relevant results in production. In particular, we concentrate on the vowel quality, since it is the focus of the perceptual experiment, and on the repetition task, since it does not involve orthography and thus provides a clearer insight into the processes at work in auditory perception than the second task we use in the production part, i.e. the reading task. Two hypotheses are examined here. The first one is based on the results we obtained in production: given the lack of significant difference of vowel quality achievement in the repetition task and following the assumption that perception precedes production, our primary hypothesis was that there would be no significant difference in terms of perceptual discrimination ability between the two contrasts /*ã*/-/*õ*/ and /*ã*/-/*ẽ*/.

However, if we consider phonetic and phonological similarities between French and Japanese [3, 15, 16], we could expect different results for the following two reasons: first, given the ongoing phonetic change in the standard French nasal vocalic system [13], which tends to weaken the difference between the realizations of /*ã*/ and /*ẽ*/, we could expect more perceptual confusions for non-native listeners between /*ã*/ and /*ẽ*/ (rather than

between /*ã*/ and /*õ*/); second, when we take into account the adaptation processes at work in the Japanese loanword phonological system, whereby /*ã* *õ* *ẽ*/ are often respectively adapted as /aN oN aN/ (e.g. ‘pain’ /p<sup>ẽ</sup>/ adapted as /pa.N/), we can see that the correspondence between the French vowel /*ẽ*/ and the Japanese bimoraic sequence /aN/ strengthens the potential confusion between the two French vowels /*ã*/ and /*ẽ*/.

Thus we expect the vowel /*õ*/ to be better perceived. Therefore our secondary hypothesis was that the discrimination rates of the /*ã*/-/*ẽ*/ contrast would be lower than the ones of the /*ã*/-/*õ*/ contrast. In order to test these two alternative hypotheses and to compare the performances of the learners in the production and the perception tasks at the same learning stage, our subjects had to carry out an AXB discrimination task.

## 2. METHOD

### 2.1. Participants

Participants were 22 Japanese university students (4 male and 22 female, mean age 19) who had been studying French as beginners for 4 months at Tokyo University of Foreign Studies at the time of the recording.

### 2.2. Material

12 monosyllabic words containing a nasal vowel: 6 with /*ã*/, 3 with /*ẽ*/ and 3 with /*õ*/, distributed in 6 pairs of words (for the /*ã*/-/*õ*/ contrast: *anse-once*, *panse-ponce*, *pan-pont* and for the /*ã*/-/*ẽ*/ contrast: *Andes-Inde*, *tante-teinte*, *tant-teint*). In order to ensure the comparability of the perception and the production tests, these same words were used in the AXB task first, and then in the production protocol (in a repetition and in a reading task). For the AXB task, the stimuli were produced by 3 native speakers (1 male and 2 female) in order to differentiate the phonetic level (discrimination can be simply based on acoustic cues) from the phonological one (subjects must rely on phonemic categories) [20]. Triplets were created with an ISI of 1s, with the target word in the middle position (e.g. *anse-once-once*) produced by the same female voice for A and B, and by a similar, a different female or a male voice for X. The final stimulus set consisted of 72 triplets (4 triplets x 6 pairs of words x 3 voice conditions), divided into two sets (36 stimuli for the /*ã*/-/*õ*/ contrast and 36 for the /*ã*/-/*ẽ*/ contrast).

### 2.3. Procedure

The participants performed one AXB discrimination task via Labguistic, a dedicated web platform, [www.labguistic.com](http://www.labguistic.com), [17]). The subjects had 5 seconds to decide whether the stimulus (in the middle) was similar to the first or to the last member of the triplet by clicking on the appropriate box (1 if the stimulus is similar to the first member of the triplet or 3 if it is similar to the last member of the triplet). The order of presentation of the stimuli was randomized for each participant.

## 3. RESULTS

The results of the AXB task were analyzed in terms of correct discrimination rates as a function of the voice condition (same voice, different female voice, male voice,  $n = 528$  for each condition) and of the contrast ( $/\tilde{a}/$  vs.  $/\tilde{o}/$  and  $/\tilde{a}/$  vs.  $/\tilde{e}/$ ,  $n = 792$  for each contrast). Statistical analyses were conducted using mixed-effects regression models [1], in which the participants and stimuli were entered as random terms.

### 3.1. Global results

The global analysis shows an effect of condition ( $F(1, 1556) = 53.86$ ,  $p < 0.001$ ), an effect of contrast ( $F(2, 1556) = 30.58$ ,  $p < 0.001$ ), but no interaction between these two variables ( $F(2, 1556) = 2.30$ , ns), indicating that the voice condition effect is identical for both contrasts.

### 3.2. Voice condition

For the voice condition, results show a significant higher discrimination rate when all tokens of the triplets are produced with the same voice (93.18%), than when they are produced with different voices (other female voice: 78.98%; male voice: 76.52%,  $p < 0.001$  in both cases), with no difference between the other female voice and the male voice. As it has been underlined in other studies [15], the gradation between the ‘same’ voice condition and each of the ‘different’ voice conditions can be attributed to the difficulty for the learners to perceptually normalize the vowels as phonological categories of the L2, taking into account inter-speakers variation. In the ‘same’ voice condition, they can thus simply rely on their ability to discriminate any pairs of sounds based on their acoustic differences, whereas a correct discrimination in the ‘different’ voice condition can only be carried out if learners have established phonemic categories in the L2. The gradation we found in our results shows that even though they are beginner learners of French with

only 4 months of French learning experience, they already have established phonemic categories. As, on the one hand, we made sure that the learners were able to process linguistic (phonemic) units rather than physical (acoustic) units and as, on the other hand, the global analysis showed no interaction between the voice condition and the contrast, we decided not to take into account the voice condition variable for the following analyses.

### 3.3. Contrast effect

For the contrast effect, our results show a significant better discrimination rate for  $/\tilde{a}/$ - $/\tilde{e}/$  than for  $/\tilde{a}/$ - $/\tilde{o}/$ , with respectively 94.19% vs. 71.59%. Taking into account the important difference in the discrimination rates between the two contrasts (22.60%), we decided to examine the individual vowel discrimination rates separately for  $/\tilde{a}/$ - $/\tilde{e}/$  and for  $/\tilde{a}/$ - $/\tilde{o}/$ . When subjects have to discriminate  $/\tilde{a}/$  and  $/\tilde{o}/$ , although the mean discrimination rate is slightly higher for  $/\tilde{o}/$  (74.75%) than for  $/\tilde{a}/$  (68.43%), results show no significant difference between both vowels ( $F(1, 772) = 1.00$ , ns). By contrast, when learners have to distinguish  $/\tilde{a}/$  and  $/\tilde{e}/$ , the discrimination rate is significantly lower for  $/\tilde{e}/$  than for  $/\tilde{a}/$ , with respectively 92.68% vs. 95.70% ( $F(1, 778) = 13.81$ ,  $p < 0.001$ ).

## 4. DISCUSSION

The aim of the study reported here was to assess the perceptual ability of beginner Japanese learners of French to discriminate the nasal vowels  $/\tilde{a}/$   $\tilde{o}$   $\tilde{e}/$  through an AXB discrimination task testing two contrasts ( $/\tilde{a}/$ - $/\tilde{o}/$  and  $/\tilde{a}/$ - $/\tilde{e}/$ ) and compare the results with their productions of the same three nasal vowels in a repetition task in the framework of a perception-production longitudinal study. Regarding the AXB study, the main results are twofold: a) in terms of contrast,  $/\tilde{a}/$ - $/\tilde{e}/$  is significantly better discriminated than  $/\tilde{a}/$ - $/\tilde{o}/$ ; b) in terms of discrimination rates, whereas no significant difference is found between  $/\tilde{o}/$  and  $/\tilde{a}/$ ,  $/\tilde{e}/$  obtains a significantly lower rate when it is contrasted with  $/\tilde{a}/$ .

Regarding the first result, since the discrimination rate is significantly higher for  $/\tilde{a}/$ - $/\tilde{e}/$  than for  $/\tilde{a}/$ - $/\tilde{o}/$ , our two initial hypotheses (no significant performance distinction between  $/\tilde{a}/$ - $/\tilde{o}/$  and  $/\tilde{a}/$ - $/\tilde{e}/$  or better discrimination for the  $/\tilde{a}/$ - $/\tilde{o}/$  contrast) proved to be wrong. Given that in our

production data from the repetition task, none of the three nasal vowels was found to be significantly better produced than the others in terms of vowel quality, we can interpret the difference we found in the AXB task as an evidence of a differential development between the interphonological categorical system in production (no difference at this stage between the three vowels in terms of targetlikeness) and in perception (considering the significant difference between the two pairs, /*ã*/-/*õ*/ and /*ã*/-/*ẽ*/). This supports the idea that the perception of the contrast precedes its actual production.

Regarding the second result (lower discrimination rate for /*ẽ*/ than for /*ã*/, whereas there is no significant difference between /*ã*/ and /*õ*/), this discrepancy can be considered in light of one of the main hypotheses of the SLM [10, 11], in which ‘new’ phones are easier to acquire than ‘similar’ phones. We can interpret this result by assuming that, for Japanese learners of French, [*ẽ*] is indeed a ‘new’ phone (as opposed to [*ã*] and [*õ*], which are phonetically ‘similar’ to [aŋ] and [oŋ]), perceptually easier to discriminate, but at the same time as difficult to produce as the two others (no significant distinction in production in the repetition task) or even more difficult (lower scores in the reading task in our production protocol). This result eventually also supports the perception-before-production principle.

## 5. CONCLUSION

The study presented here is embedded in a longitudinal perception-production study of the acquisition of French phonology by beginner Japanese university students in Japan. In previous studies carried out within our project regarding the production of nasal vowels by advanced Japanese learners of French [7, 18], a quality ranking in terms of rates of target-likeness had emerged: /*õ*/ was better produced than /*ã*/ which was better produced than /*ẽ*/.

This ranking was convergent with the ones obtained in different studies with English learners of French ([2, 14]). When we contrasted these results of advanced learners with those obtained among beginner learners, we discovered a different pattern of production: while both populations exhibited rather high performances in terms of nasality achievement (nasal vs. oral vowels), the vowel quality assessments differed. Indeed, if we only consider the results of the repetition task for the sake of comparison with the perception experiment, as

explained above, it appears that the quality assessments of the vowels produced by the beginner learners show no significant difference among the three vowels. This gap in the beginner-advanced learners results led us to hypothesize that novice learners were able to acquire the inter-category contrast between oral and nasal vowels faster than the finer-grained intra-category distinction between the three nasal vowels /*ã*/, /*õ*/, /*ẽ*/. One possible developmental path in production could then be as follows: 1) nasal (no distinction between the three nasal vowels) vs. oral; 2) /*õ*/ vs. /*ã*/ vs. /*ẽ*/. On the other hand, the results we elicited in perception point to an early differentiation between at least the two contrasts tested here, which can then be interpreted as a finer intra-category discrimination ability in perception (vs. the absence of such ability in production at this stage). Finally, if we take into account not only the results of the repetition task, but also those of the reading task, the emerging pattern (lower score for /*ẽ*/ but no significant difference between /*õ*/ and /*ã*/), we could further hypothesize an intermediary stage by which the intra-categorical distinction between the three vowels first proceeds through a binary distinction between a merger /*õ* ~ *ã*/ category vs. a single /*ẽ*/ category. This would fit with the results obtained in the AXB task, with lower rates of discrimination between /*õ*/ and /*ã*/ because of the perceived similarity between the two) and higher rates of discrimination between /*ẽ*/ and /*ã*/ because of the ‘novelty’ of the [*ẽ*] phone and its perceived dissimilarity from the more ‘similar’ [*ã*] phone.

In any case, the comparison of the results obtained in perception and in production point to a possible differential development with an intra-categorical distinction between the three nasal vowels more quickly established in perception, thus supporting the view according to which perception precedes production in L2 phonology acquisition. Our developmental hypothesis remains to be fully tested over the course of the two years of our longitudinal study (with 4 sessions, each after 6 months of further learning), which should be possible once our data are fully analyzed. This will also allow us to check whether the later stage of development of the nasal vowels system in the interphonology of the learners corresponds to our previous assumptions, i.e. the /*õ*/ > /*ã*/ > /*ẽ*/ ranking found in previous studies among advanced learners.

## 6. REFERENCES

- [1] Baayen, R. H., Davidson, D. J., Bates, D. M. 2008. Mixed effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390-412.
- [2] Charlton, A. 2012. French nasal vowels: acquisition by advanced English learners. *SLP*, York. (poster presentation).
- [3] Detey, S. 2005. *Interphonologie et représentations orthographiques. Du rôle de l'écrit dans l'enseignement/apprentissage du français oral chez des étudiants japonais*. Unpublished PhD Dissertation, University of Toulouse.
- [4] Detey, S., Racine, I., Eychenne, J., Kawaguchi, Y. 2014. Corpus-based L2 phonological data and semi-automatic perceptual analysis: the case of nasal vowels produced by beginner Japanese learners of French. *Proc. 15<sup>th</sup> Interspeech* Singapore, 539-544.
- [5] Detey, S., Racine, I., Kawaguchi, Y. 2014. Des modèles prescriptifs à la variabilité des performances non-natives : les voyelles nasales des apprenants japonais et espagnols dans le projet IPFC. In: Durand, J., Kristoffersen, G., Laks, B. (eds), *La phonologie du français : normes, périphéries, modélisation. Hommage à Chantal Lyche*. Paris: Presses Universitaires de Paris Ouest, 197-226.
- [6] Detey, S., Racine, I., Kawaguchi, Y., Zay, F. in press. Variation among non-native speakers: the InterPhonology of Contemporary French. In: Detey, S., Durand, J., Laks, B., Lyche, C. (eds), *Varieties of Spoken French: a source book*. Oxford: Oxford University Press.
- [7] Detey, S., Racine, I., Kawaguchi, Y., Zay, F., Buehler, N., Schwab, S. 2010. Evaluation des voyelles nasales en français L2 en production : de la nécessité d'un corpus multitâches. *Proc. CMLF 2010* Paris, 1289-1301.
- [8] Durand, J., Gut, U., Kristoffersen, G. (eds) 2014. *The Oxford Handbook of Corpus Phonology*. Oxford: Oxford University Press.
- [9] Escudero, P. 2006. Second language phonology: the role of perception. In: Pennington, M. (ed.), *Phonology in Context*. New York: Palgrave MacMillan, 109-134.
- [10] Flege, J. E. 1987. The production of 'new' and 'similar' phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, 47-65.
- [11] Flege, J. E. 2003. Assessing constraints on second-language perception and production. In: Meyer, A., Schiller, N. (eds), *Phonetics and Phonology in Language Comprehension and Production, Differences and Similarities*. Berlin: Mouton de Gruyter, 319-355.
- [12] Gut, U. 2009. *Non-native Speech: a Corpus-based Analysis of Phonological and Phonetic Properties of L2 English and German*. Bern: Peter Lang.
- [13] Hansen, A. B. 2001. Les changements actuels des voyelles nasales du français parisien : confusions ou changement en chaîne ? *La linguistique*, 37(2), 33-48.
- [14] Inceoglu, S. 2014. Effect of multimodal training on the perception of French nasal vowels. *Concordia Working Papers in Applied linguistics*, 5, 311-321.
- [15] Kamiyama, T. 2009. *Apprentissage phonétique des voyelles du français langue étrangère chez des apprenants japonophones*. Unpublished PhD Dissertation, University of Paris La Sorbonne Nouvelle.
- [16] Labrune, L. 2012. *The Phonology of Japanese*. Oxford: Oxford University Press.
- [17] Ménétrey, P., Schwab, S. 2014. *Labguistic*: a web platform to design and run speech perception experiments. In: Congosto Martín, Y., Montero Curiel, M.L., Salvador Plans, A. (eds), *Fonética Experimental, Espacio Europeo, Educación Superior e Investigación: Fonética y Fonología (vol I)*. Universidad de Cáceres y Universidad de Sevilla, 543-556.
- [18] Racine, I., Detey, S., Buehler, N., Schwab, S., Zay, F., Kawaguchi, Y. 2010. The production of French nasal vowels by advanced Japanese and Spanish learners of French: a corpus-based evaluation study. *Proc. 6<sup>th</sup> New Sounds* Poznan, 367-372.
- [19] Racine, I., Detey, S., Zay, F., Kawaguchi, Y. 2012. Des atouts d'un corpus multitâches pour l'étude de la phonologie en L2: l'exemple du projet Interphonologie du Français Contemporain (IPFC). In: Kamber, A., Skupien, C. (eds), *Recherches récentes en FLE*. Bern: Peter Lang, 1-19.
- [20] Strange, W., Shafer, W. L. 2008. Speech perception in second language learners: The re-education of selective perception. In: Hansen Edwards, J. G., Zampini, M. L. (eds), *Phonology and Second Language Acquisition*. Amsterdam/Philadelphia: John Benjamins, 153-192.

---

<sup>1</sup> This research has been supported by the Japanese Society for the Promotion of Science through a Grant-in-Aid for Scientific Research (B) n°23320121 to Sylvain Detey. We wish to thank Yuji Kawaguchi, Mariko Kondo, Jacques Durand, Marie-Laure Sandoz, Marion Didelot, Mito Matsuzawa, Tsuyoshi Umeno, Kahori Ohmura, as well as all the students who participated in the study. A first version of the work presented here benefited from the comments of the audience at CIL19 in Geneva, we would like to thank them as well.