

Towards an understanding of the cognitive processes involved in nonnative speech sounds processing

Difficulties encountered when perceiving and pronouncing nonnative contrasts have been widely documented in the Second Language Acquisition (SLA) literature. Some models were proposed in order to predict the success of the nonnative speech perception depending on the phonetic similarity between the nonnative and the native categories (Best, 1995 ; Flege, 1995). Furthermore, a great amount of factors contributing to the success of perception-production of a second language (L2) have been reported (Piske et al., 2001; Strange & Shafer, 2008).

However, while the variables improving nonnative sounds processing have been under thorough investigation in SLA, few experiments have addressed the *underlying cognitive processes*. Failures in perceiving nonnative speech sounds are commonly accounted for by reference to the notion of interference of the first language (L1) into the L2 (Troubetzkoy, 1986; Flege, 1995; Kuhl et al., 2008; Strange, 2011;). Indeed, exposure to the native language in infancy results in allocating more attentional weight to L1-relevant phonetic cues than to irrelevant ones. Therefore, following 'attention-to-dimension' models (Iverson & Kuhl, 1995; Francis & Nusbaum, 2002; Nosofsky, 1986), phonetic learning in L2 may be considered as the result of two mechanisms of selective attention: withdrawal of attention from irrelevant phonetic dimensions and enhancement of attention on the relevant ones. This is in agreement with evidence showing that new phonological contrasts may be acquired after training redirecting attention to the relevant phonetic cues (Lively et al., 1994; Francis & Nusbaum, 2002).

While selective attention mechanisms have been further investigated (Kondaurova & Francis, 2010), few research has addressed the role of *memory processes*. However, as suggested by the inverse relationship between the duration of the inter-stimulus interval (ISI) and discrimination performance (Pisoni, 1973; Werker & Tees, 1984), different kinds of memory processes are involved in speech sounds processing. Specifically, it has been claimed that some parameters such as the duration of the inter-stimulus interval (ISI) (Pisoni, 1973), or the task (Gerrits, 2001), influence discrimination performance by eliciting a differential reliance on auditory short-term memory. In fact, when discriminating two sounds the performance strongly depends on the availability of the memory trace because comparison between a preceding sound and an incoming sound relies on the use of fresh information (Näätänen, 2011). Gerrits (2001) suggests that the decay of the sound traces after a few tens of milliseconds prompts the subject to rely on the phonological representations stored in long-term memory.

As very few experiments have attempted to shed light on the cognitive functioning related to nonnative speech processing, our work is meant as a contribution to this topic by addressing the attention and memory mechanisms underlying the modification of perception and production modes. Indeed, the automatized perception routines developed by the individual to process the native language usually turn out to be inefficient in processing nonnative speech sounds so that the nonnative learner faces an effortful, attention-demanding task consisting in overcoming the influence of usual routines and

developing new ones (Strange, 2011). Our project addresses the general issues of how the different memory systems (auditory sensory memory, short-term memory, long-term representations) are involved when processing nonnative sounds and how attentional mechanisms of inhibition and enhancement contribute to this operation. More specifically, in this paper we aim to investigate the issue of how different memory systems are activated or inhibited by the learner while processing native and nonnative speech sounds. We propose here a first exploratory study in which we examine the variations in discrimination performances under various conditions that have been designed to (dis)favour the recruitment of short-term vs. long-term memory processes. To this end, three variables likely to influence the processing mode adopted by the subject are manipulated: first, the type of stimuli making the discrimination pair; second the inter-stimulus interval; third, the instructions given to the subject prior to the task.

The experimental design is as follows. Adult French-speaking subjects whose experience with foreign languages is limited (mimicking a beginning learner of L2) are administered pairs of vowels in a same/different discrimination task. Stimuli are eight synthetic vowels based on the /y-u/ French contrasts which acoustic parameters are carefully controlled. They are labelled as follows (see schematic representation in Fig.1): prototypical /y/ (PrY), prototypical /u/ (PrU), acceptable exemplar of /y/ (AcY), acceptable exemplar of /u/ (AcU), poor exemplar of /y/ (PoY), poor exemplar of /u/ (PoU), nonnative sound out of the /y/ domain (NnY), nonnative sound out of the /u/ domain (NnU). A total of six pairs are to be discriminated, i.e. PrY-PrU, AcY-AcU, PoY-PoU, NnY-NnU, PrY-PoU, AcY-NnU. They are presented in separate blocks of 30 trials including the two orders of presentation of the 'different' pair and the two 'same' pairs (e.g. for the PrY-PrU pair: PrY-PrU, PrU-PrY, PrY-PrY, PrU-PrU). Each pair is administered in eight conditions from the declination of two additional variables, i.e. ISI (very short:400ms; short:800ms; long:1200ms; very long:1600ms) and the instructions given to the subject prior to the task, i.e. first "discriminate these two speech sounds: say 'different' as soon as you detect any difference, even a tiny one", then "discriminate these two sounds that stem from an attempt of the experimenters to synthesize the French vowels /y/ and /u/; say 'different' when you feel like one sound is an /y/-like sound and the other one is an /u/-like sound, even when the quality is very poor".

Our working hypotheses are the following: (i) the longer the ISI, the less the subject will be able to rely on sensory traces to make a decision, which should particularly effect performances on pairs including nonnative sounds; (ii) the more prototypical of a native category is a stimulus, the less the subject will be able to inhibit long-term memory processes, which may decrease performance under (very) short ISI; (iii) the first type of instructions should favour acoustico-phonetic processes based on sensory traces and short-term memory, whereas the second type of instructions should put the subject in a 'phonological' mode, which in turn would affect discrimination performances.

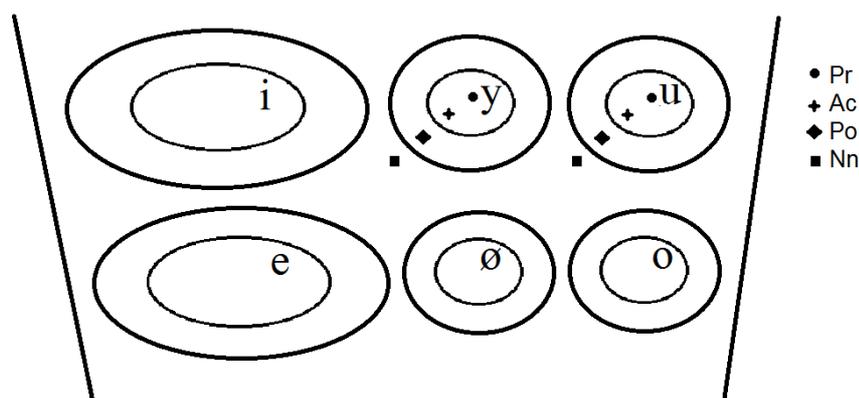


Fig.1 Schematic representation of the position of the eight stimuli in the French vowel space

References

- 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 Research, Baltimore: York Press, pp. 171-203.
- Cowan, N., Morse, P.A. (1986). The use of auditory and phonetic memory in vowel discrimination. *Journal of Acoustical Society of America*, 79 (2), 500-507.
- Flege, J. (1995). Second-language Speech Learning: Theory, Findings, and Problems. In W. Strange (Ed) Speech Perception and Linguistic Experience: Issues in Cross-language research, Baltimore : York Press, pp. 229-273.
- Francis, A.L., & Nusbaum, H.C. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance*, 28(2), 349-366.
- Gerrits, E. (2001). The categorization of speech sounds by adults and children. Unpublished doctoral dissertation, Utrecht University.
- Iverson, P, & Kuhl, P.K. (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *JASA*, 97 (1), 553-562.
- Kondaurova, M.V. & Francis, A.L. (2010). The role of selective attention in the acquisition of English tense and lax vowels by native Spanish listeners: comparison of three training methods. *Journal of Phonetics*, 38(4), 569-587.
- 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*, 363, 979-1000.
- Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/: III. Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America*, 96, 2076–2087.
- Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, 38, 1-21.
- Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *J. Exp. Psychol. Gen.*, 115, 39-57.
- Piske, T., MacKay, I. & Flege, J. (2001). Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, 29, 191–215.
- Pisoni, D.B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics*, 13 (2), 253-260.
- Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech : A working model. *Journal of Phonetics*, 39, 456-466.
- Strange, W., & Shafer, V. L. (2008). Speech perception in second language learners: The re-education of selective perception. In J. G. Hanson Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 193-218). Amsterdam: John Benjamins.
- Troubetzkoy, N.S. (1986) *Principes de Phonologie*, Paris : Klincksieck.
- Werker, J.F., Tees, R.C.(1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of Acoustical Society of America*, 75 (6), 1866-1878.