HAL
open science

Role of prosody in segmental perception in L2

Takeki Kamiyama, Shigeko Shinohara

To cite this version:

Takeki Kamiyama, Shigeko Shinohara. Role of prosody in segmental perception in L2. Sophia University Working Papers in Phonetics 2010, 2010, pp.10-18. hal-00587963

HAL Id: hal-00587963

https://hal.science/hal-00587963

Submitted on 22 Apr 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ROLE OF PROSODY IN SEGMENTAL PERCEPTION IN L2

KAMIYAMA, Takeki[1], *SHINOHARA, Shigeko*[2]

1. Laboratoire de Phonétique et de Phonologie (UMR7018), CNRS / Sorbonne Nouvelle, Paris
2. Phonetics Laboratory, Sophia University, Tokyo / CNRS

ABSTRACT

This paper focuses on the role of prosody in segmental perception of CCV cluster vs. CVCV in French by Japanese speakers learning French. We studied Japanese speakers' perception of minimal pair words in French of the type /blɛt/ vs. /bulɛt/ embedded in different intonation contours, with higher or lower F0 on the cluster. Results indicate that tonal prominence facilitates the perception of segmental sequences in L2.

Keywords: Prosody, L2 acquisition, Japanese, French

1. INTRODUCTION

In the very early stage of L1 acquisition, infants seem to have the predisposition to be sensitive to the prosodic cues to phrase boundary. Previous studies report a heavy reliance on the stress patterns for word segmentation in the early stage of lexical acquisition (newborn to 7.5month) ([2] for French; [7] for English), which moves towards multiple strategies within a few months [7]. In the production of young children, consonant clusters are mostly resolved through consonant deletion [1]: the result is that prosodic patterns remain more or less intact as compared to the adult target. In contrast, it is shown in recent loanword studies that consonant clusters are overwhelmingly resolved through vowel epenthesis rather than through consonant deletion [8][10]. This pattern seems to be also reflected in adult production of L2 for Japanese speakers [9]. A consequence of vowel epenthesis is a distortion of prosody in the target language, the number of nuclei being increased. We wonder if this production pattern reflects adult speakers' insensitivity toward prosodic information in L2. In this study we would like to test if prosodic information is still exploited in access to segmental information in L2 speakers of French.

2. BACKGROUD

In French, the fall in fundamental frequency (F0) at the end of affirmative sentences tends to extend to the last three syllables. [11] and [12] characterize the tonal patterns of affirmative and yes-no question sentences as follows. If the last lexical word contains less than three syllables and is preceded by a function word, which tends to be pronounced low, the beginning of that lexical word can have an optional local pitch peak. In the case of total questions, F0 stays relatively flat until the last syllable, where a sharp rise occurs. Following these observations, given that the first syllable of disyllabic words can be produced with a higher pitch than the following one in affirmatives and not in total questions, we hypothesize that the higher pitch, presumably more prominent prosodic patterns, in affirmatives facilitate the perception of segmental sequences.

We undertook a series of experiments. We chose minimal pair words in French differing only by the presence or absence of a vowel, and which are adapted into Japanese as words containing the same segments, but one requires a vowel epenthesis because Japanese does not allow complex onset (e.g. *boulette* /bulɛt/ 'small ball' > /buretto/, and *blette* /blɛt/ 'beet' > /b*u*retto/; the epenthetic vowel is marked by italics) [9]. As to the perception of clusters, Japanese speakers are known to perceive an illusory epenthetic vowel (e.g. /ebzo/ perceived as /eb*u*zo/ [6]). Therefore, we can suppose that they have difficulties differentiating these minimal pairs. However, this task might be facilitated by the prosodic prominence in affirmatives. In Experiment 1, we recorded the minimal pair words in the two types of intonation in order to confirm the reported patterns, and in Experiments 2 and 3, these recordings were used for auditory perception tests with native speakers of Japanese learning French as a second language.

3. EXPERIMENT 1: PRODUCTION BY NATIVE SPEAKERS OF FRENCH

3.1. Methods

The corpus consists of 28 words (14 pairs) of CCV- and CVCV- types embedded in 68 sentences. These pair words include: *briques* /bʀik/ vs. *bourriques* /buʀik/; *glue* /gly/ vs. *goulue* /guly/; *plaine* /plɛn/ vs. *poulaine* /pulɛn/ etc... The sentences were integrated into two different contexts that would elicit affirmative and question intonations such as *Je te dis de quoi il s'agit./ De quoi s'agit-il ? Il s'agit de .../ ?* ('I'll tell you what it's about./ What is it about? It's about …Is it about …?').

2 female native speakers of French read the sentences 3 times, which were recorded on a Minidisc recorder in a semi sound-proof room and converted to WAVE files with a sampling rate of 22050 Hz, 16 bits.

3.2. Analysis

F0 was measured on Praat first to observe the general patterns, secondly for a more systematic comparison between the mono- and disyllabic words by normalizing F0 on the time axis.

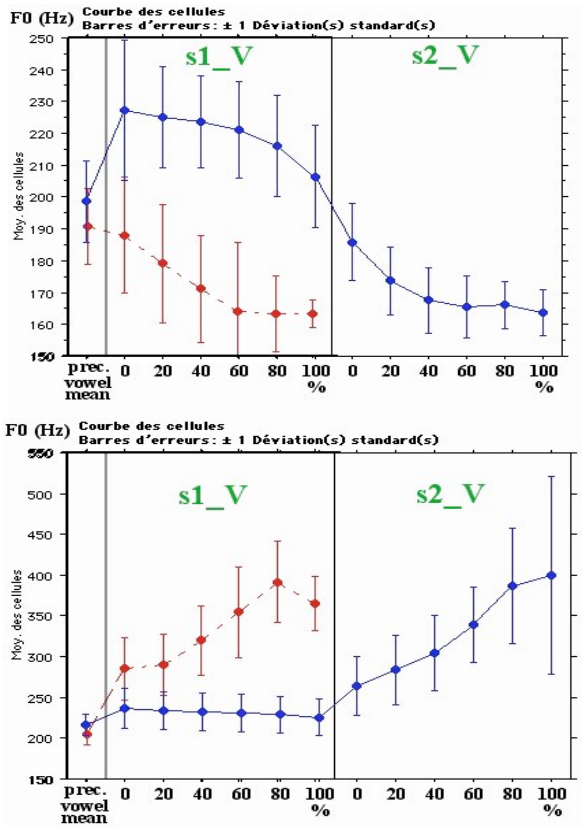
3.3. Results

The 2 speakers showed similar F0 patterns in the interrogatives, but not in the affirmatives in particular with disyllabic words.

Concerning the affirmatives, Speaker 1 tended to create a peak during the first syllable of the disyllabic target words before a fall, which was not observed in monosyllabic words, while Speaker 2 did not realize this peak clearly. As for the questions, both speakers showed a flat F0 pattern followed by a sharp rise during the last syllable of the sentence.

Figure 1 shows the F0 patterns of 32 sentences produced by Speaker 1. F0 (mean) during the vowel preceding the target word does not vary much from mono- to di- syllabic words (190 and 199 Hz for the affirmatives, 205 and 217 Hz for the questions). Concerning the target word in the affirmatives, it rises to 227 Hz at the beginning of the first vowel, and then falls to 186 Hz at the beginning of the second one, which is quite similar to that of the first vowel in monosyllabic words (188 Hz). As for the interrogatives, F0 rises at the beginning of monosyllabic words (285 Hz) and at the beginning of the second vowel of disyllabic words (264 Hz), while it stays flat (slightly falling) during the first syllable of disyllabic words.

Figure 1: Mean F0 of the affirmatives (left) and the interrogatives (right) with mono- (dashed lines) and disyllabic words (plain lines) pronounced by Speaker 1 (16 sentences each). The first value ("prec. vowel mean") indicates the mean F0 of the vowel preceding the target word. The time axis is normalized, i.e. 0% and 100% points correspond to the beginning and the end of the vowel. Error bar: 1 standard deviation (SD).



To sum up, these findings seem to confirm, overall, the expected tendencies: 1) the questions show a flat F0 pattern followed by a sharp rise during the last syllable of the sentence; 2) for the affirmatives, only one of the two speakers had expected F0 rise on the first syllable. Thus, besides the expected patterns, speaker difference was also observed.

4. EXPERIMENT 2: PERCEPTION BY JAPANESE-SPEAKING LEARNERS I

In our second experiment, native speakers of Japanese learning French as a second language listened to the sentences and tried to identify the target word. The aim is to examine the Japanese learners' ability to distinguish the two types of

words (CCV and CVCV) in relation to the prosodic pattern of the sentence.

4.1. Methods

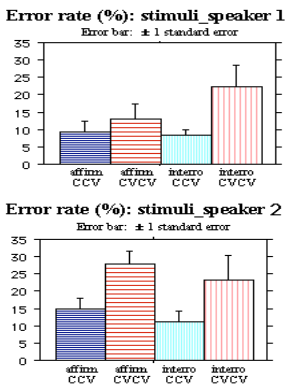
We used 48 out of the 68 sentences read by the 2 speakers in Experiment 1. The task was to listen to a sentence and decide which of the minimal pair words they had heard. The words were presented in French orthography on an answer sheet.

9 Japanese speakers living in Paris participated. They had lived there for 2 to 8 months at the time of the experiment. They were instructed to circle the word and the confidence rating corresponding to their perception. The subjects listened to the list of stimulus sentences pronounced by Speaker 1 once, and to those by Speaker 2 once after a break.

4.2. Results

The results are shown in Figure 2. First, the overall error rate was higher for the stimuli pronounced by Speaker 2. Recall that she did not make a clear peak in the affirmatives with disyllabic words. Secondly, there was little difference between the interrogatives produced by the two speakers, while the affirmatives lead to a higher error rate for Speaker 2 (significant difference for CVCV: t_{22} = 2.53, p = 0.02). Thirdly, the error rate is higher for CVCV words (significant effect for both speakers: $F_{(1, 44)}$ = 4.38, p = 0.04, $F_{(1, 44)}$ = 7.38, p = 0.01, respectively), and this tendency is particularly clear for Speaker 1's interrogatives.

Figure 2: Error rate for stimuli pronounced by Speakers 1 (left) and 2 (right). "affirm" stands for affirmatives, "interro" for interrogatives, "CCV" for CCV words, and CVCV for CVCV words (9 listeners x 48 stimuli for each speaker).

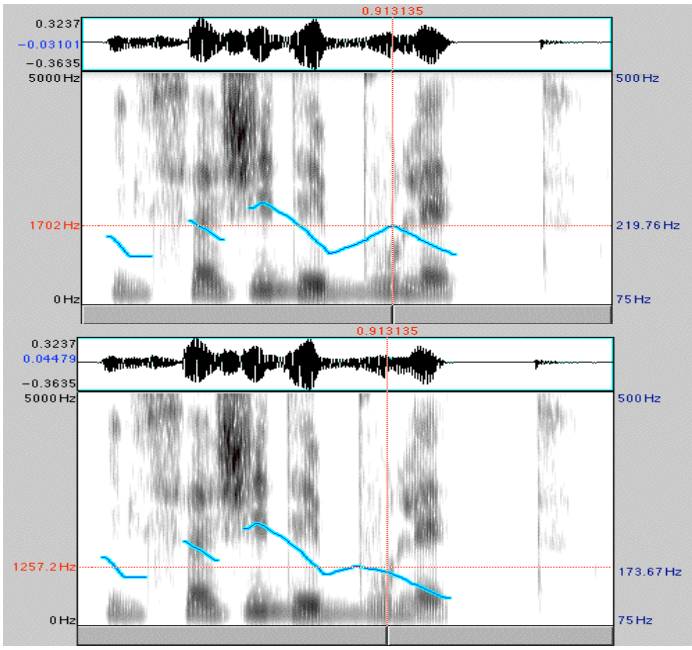


These results suggest the following: when the test sentences are better articulated, the difference between the affirmatives and the interrogatives is clearly obtained. However, it is not clear if the tendency is entirely due to the prosodic differences, or also to some other factors such as hypo/hyper articulation.

5. EXPERIMENT 3: PERCEPTION BY JAPANESE-SPEAKING LEARNERS II

For this reason, we conducted another experiment with the same listeners, using stimulus sentences differing solely in f0 patterns of the last word.

Figure 3: Sentence "*Il s'agit des boulettes.*" pronounced by Speaker 1 and resynthesized with a simplified F0 curve (left); with a lowered f0 on the syllable /bu/ in the word /bulɛt/ (right).



5.1. Methods

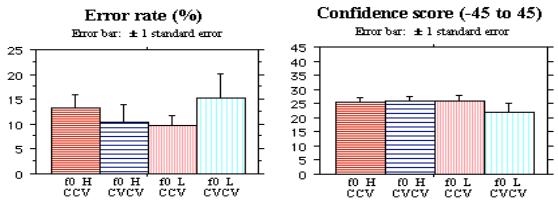
16 affirmative sentences (8 minimal pairs) pronounced by Speaker 1 were re-synthesized with a simplified f0 curve on Praat. Then f0 was manipulated so that each sentence would have two different f0 patterns during the last word, namely, one with a peak higher than the preceding function word, the other with a lower peak resulting in a continuous fall (see Figure 3).

The procedures were similar to those of Experiment 2, except that the list of stimuli was presented twice in different orders.

5.2. Results

Figure 4 shows the error rate and a score calculated on the basis of errors and confidence rating (if all subjects answered correctly with a top confidence rating, the score will be 5 x 9 listeners = 45).

Figure 4: Error rate (left) and score calculated on the basis of errors and confidence rating (right: - 45 to 45). 9 listeners x 32 stimuli x 2 repetitions.



Even though there is no significant difference, CVCV words tend to be difficult to identify if pronounced with a low f0 pattern. This tendency corresponds to the difference found between the stimuli of the two speakers in Experiment 2.

To sum up the results of these two perceptual experiments, the tonal salience in the affirmatives seem to aid, as expected, the Japanese learners to differentiate the two types of segment sequences.

6. DISCUSSION

The findings in Experiment 1 indicate that the first syllable of disyllabic words preceded by a function word at the end of affirmative sentences tends to be pronounced with a higher F0 than that of the beginning of monosyllabic words. However, the realization of this tendency seems to depend on the strategy and the style adopted by the speaker: it was observed clearly in the recordings of Speaker 1, but not in those of Speaker 2. Speaker 1, who is a professional speech therapist, tended to hyperarticulate, while Speaker 2 did not.

The results in Experiment 2 could be explained better by taking into account this inter-speaker difference reflected in the stimuli. Concerning the higher error rate for CVCV words in questions produced by Speaker 1, we could suppose that CVCV tends to be perceived as CCV in the questions more than in the affirmatives. This is probably because the first syllable in the CVCV words in the total questions, which tends to be pronounced with little f0 movement approximately at the same f0 range as the preceding function word, is less easily perceived as a distinct syllable than that in affirmatives. Thus, the F0 rise in affirmative sentences might have facilitated the perception of the first syllable of the disyllabic words.

However, as for the sentences produced by Speaker 2, the error rate of CVCV syllables in the affirmatives is almost as high as the interrogatives, suggesting that the first syllable in CVCV words in affirmatives, which is prosodically less prominent for Speaker 2 than for Speaker 1, did not help the listeners perceive this as a distinct syllable.

What is curious about the results is that the overall error rate is higher for CVCV words than for CCV words, implying that the learners perceived a cluster where there was a vowel between the two consonants more often than they perceived an inserted vowel where there was a cluster. This might be due to the learner's perceptual hypercorrection, but it seems to contradict earlier findings in which Japanese listeners perceived an inserted vowel in clusters [6]. Further investigation is needed to clarify this point.

Regarding the final curious result, another interpretation concerns the prominence related to duration. Segmental distinctiveness can be enhanced by prosodically prominent position; and there may be more than one type of prosodic prominence. Durational prominence can be independent of tonal prominence biasing the perception in another dimension. While accentuation in French is a controversial issue in French phonology [5][3], it is generally agreed that the phrase final syllables have longer duration than preceding syllables [4][11]. It might be the case that even though the first syllable of the disyllabic words have distinctive roles, the salience of those syllables are inferior to the phrase final ones so that their existence is confused with zero (=CCV sequence).

7. CONCLUDING REMARKS

The findings in the present study seem to indicate that tonal prominence facilitates the perception of segmental sequences in L2 perception, suggesting the importance of prosody in L2 phonetics.

We have discussed the role of prosody in relation to segments in L1 and L2 acquisitions. Assuming that prosody has a very important role in the early stage of L1 acquisition and some role in L2 acquisition, the next questions to ask are: 1) How much impact does the prosody have in L2 segmental acquisition; 2) If it turns out that the weight of prosody in L1 is still greater than that in L2, when and how does it decrease; in other words, how does L1 acquisition grammar weighing on prosody change to adult grammar.

8. REFERENCES

[1] Bernhardt B. Handford, Stemberger J.P. 1998. *Handbook of Phonological Development (From the Perspective of*

Constrain-Based Nonlinear Phonology), San Diego: Academic Press.

[2] Christophe, A. Dupoux, E., Bertoncini, J., Mehler, J. 1995. Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition, *Journal of the Acoustic Society of America, 3*, 1570-80.

[3] Delais-Roussarie, E. 2005. *Phonologie et grammaire : Etudes et modélisation des interfaces prosodiques*, mémoire de synthèse présenté pour l'obtention d'une Habilitation à diriger des recherches, ERSS Université de Toulouse le Mirail

[4] Delattre, P. 1939. Accent de mot et accent de groupe, *The French Review* 13 (2), 141-146.

[5] Dell, F. 1984. L'accentuation dans des phrases en français. In F. Dell, D. Hirst and J.R. Vergnaud (eds.) *Forme sonore du langage: structure des représentation en phonologie*, 65-122. Paris: Hermann.

[6] Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., Mehler, J. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25(6), 1568-1578.

[7] Jusczyk, P., Houston, W., Derek. M., Newsom, M. 1999. The beginnings of word segmentation in English-learning infants, *Cognitive Psychology*, 39, 159-207.

[8] Paradis, C., Béland, R. 2002. Syllabic Constraints and Constraint Conflicts in Loanword Adaptations, Aphasic Speech and Children's Errors, Jacques Durand et Bernard Laks (eds.), *Phonetics, Phonology and Cognition*. Oxford: Oxford Univ. Press, 191-225.

[9] Shinohara, S. 1992. *Prononciation du français par des locutrices japonaises -le cas de l'insertion de voyelles dans les groupes consonantiques avec /l/ et /R/,* DEA Diss., Univ. of Paris V.

[10] Uffmann, C. 2004. *Vowel Epenthesis in Loanword Phonology*. Ph. D. dissertation, Univ. Marburg.

[11] Vaissière, J. 2002. Cross-linguistic prosodic transcription: French versus English. *Problems and methods in experimental phonetics, In honour of the 70th anniversary of Prof. L.V. Bondarko*, N. B. Volslkaya, et al. St.-Petersburg, St.-Petersburg State Univ.: 147-164.

[12] Vaissière, J. 2006. *La Phonétique.* Paris: PUF.