Can orthography influence second language syllabic segmentation? Japanese epenthetic vowels and French consonantal clusters✧

Sylvain Detey [a*] , Jean-Luc Nespoulous [b]

[a] *Laboratoire DYALANG, Université de Rouen, IRED, 7 rue Thomas Becket 76821 Mont Saint-Aignan cedex, France*
[b] *Laboratoire Jacques Lordat, Université de Toulouse II - Le Mirail, 5 allées Antonio Machado 31058 Toulouse cedex 1, France*

Abstract

After introducing what is known about potential interactions between phonetic/phonological and orthographic representations in first (L1) and second (L2) language speech perception studies, loanword and inter- phonology, and literacy-related phonological awareness research, the paper describes the case of Japanese learners of French, with particular emphasis on the syllabic/moraic dimension of their interphonology development. We concentrate on French biconsonantal clusters of the Obstruent + Liquid (/r/ and /l/) and /s/ + Plosive type. 62 Japanese university students in Japan perform a task of syllabic segmentation of non-words presented in three conditions: auditory, visual and synchronous audiovisual. The results suggest a possible influence of orthography on L2 syllabic representations, as the audiovisual and visual conditions trigger more epentheses than the auditory condition. Six arguments are combined to account for these results: working memory, metaphonological awareness, loanword sociophonology, phonetics versus phonology, perceptual constraints and attentional resources. In light of this preliminary study, we conclude that the orthographic factor should not be neglected in L2 speech perception studies, loanword phonology and interphonology research.

Keywords: Interphonology, Orthography, Perception, French, Japanese, Syllable

[*]Corresponding author. Tel.: +33 2 35 14 69 44; Fax: +33 2 35 14 69 40.
E-mail addresses: sylvain.detey@univ-rouen.fr (S. Detey), nespoulo@univ-tlse2.fr (J.-L. Nespoulous).

1. Introduction

One of the main tasks beginner learners of a second language (L2) [1] have to perform is to segment continuous speech stream into comprehensible and memorable units, as their processing and working memory capacity are limited. In order to study L2 speech segmentation by adult learners[2], we need to consider phonological theories (phonological constraints and lexicon) (Archangeli, 1997), speech perception theories (perceptual process) (Ségui, 1997) and second language acquisition theories (influence of first language (L1) specific and universal factors on L2 perception and learning) (Eckman 1977, 2004; Broselow, 1983; Tarone, 1987; Uffmann, 2003), since the segmentation process can be influenced by phonological, phonetic and other psycholinguistic factors.

✧ We would like to thank Jacques Durand (ERSS, University of Toulouse II) for his very precious help in this work, two anonymous reviewers for their useful comments on a previous version of this article and Nicolas Ballier (LLI, University of Paris XIII) for his help with the final manuscript.

[1] "Second language" is used here as an equivalent of "foreign language". In our case, since the subjects learned English in secondary school, French could be considered a third rather than a second language for them.

[2] The situation would be different with children, since age and education would have to be more precisely taken into account. Contrary to adults, young children are still in the process of acquiring their own L1 and their ability to segment speech varies with age and education (supra-syllabic *vs.* syllabic *vs.* sub-syllabic units).

1.1. Speech perception units, loanword phonology and interphonology

Speech segmentation process both in L1 and L2 has been the focus of many psycholinguistic studies in the past[3], and the pioneering work of Mehler and his colleagues on French (Mehler, Dommergues and Frauenfelder, 1981) highlighted the role played by the syllable in this process. Since then, several studies have shown that this role differs according to the listener's L1 (Cutler, Mehler, Norris and Ségui, 1986), and that segmentation procedures vary according to the rhythmic characteristics of the L1 (Ségui and Ferrand, 2002) (syllable-based, stress-based, mora-based or others (Ramus, Dupoux and Mehler, 2003)). Due to these rhythmic, but also phonotactic and segmental characteristics, L2 input is often altered in the process, resulting in a transformed output. These transformations (especially vowel epentheses and consonantal deletions) have been a source of interest in two phonological fields in particular, namely loanword phonology (Shinohara, 1997; Jacobs and Gussenhoven, 2000) and interphonology (Broselow, 1983; Tarone, 1987). Examples of loanwords in Japanese can be found in (Kamiya, 1994): "*sâbisu*" (from "service"), "*puree*" (from "play"), "*kurabu*" (from "club"), "*sutokkingu*" (from "stokings"), etc. Recently, several researchers (Escudero, 2004; Escudero and Boersma, 2004) have been trying to combine speech perception research and phonological theories, leading to a renewed debate in loanword phonology between perceptual and phonological approaches of the transformations of L2 input into L1 output (Peperkamp and Dupoux, 2003; Smith, 2005; Rose and Demuth, 2006).

1.2. Orthography, phonological awareness and phoneme-grapheme activations

Even though it has been mentioned by some authors (for instance Flege, 1990; Pennington, 1996), the orthographic factor seems to have been rather overlooked until recently (Cross, 2002; Young-Scholten, 2002; Byung-jin, 2004; Vendelin and Peperkamp, 2005; Bassetti, 2006), in both loanword phonology and speech perception research. Although the role of literacy in phonological awareness development has been the topic of several studies in the past (Morais, Cary, Alegria and Bertelson, 1979; Morais, Content, Cary, Mehler and Ségui, 1989; Ventura, Kolinsky, Brito-Mendes and Morais, 2001)[4], this factor has been neglected in other areas probably on account of the "primacy of speech" (Durand, 2000). Yet, these studies show that orthographic experience modifies phonological awareness (Him Cheung and Chen, 2004), especially phonemic awareness (Bentin, 1992), and several results suggest that orthographic representations might play a role in speech perception (Chéreau, Hallé and Ségui, 1999; Grainger, Nguyen Van Kang and Ségui, 2001; Ziegler, Muneaux and Grainger, 2003), at least through bidirectional activations between graphemes and phonemes (Dijkstra, Frauenfelder and Schreuder, 1993). Hearing a lexical unit might activate an orthographic representation, which might in turn influence phonological judgments. Consequently, when we take into account the phonological, phonetic and orthographic differences between L1 and L2 systems of adult learners, we have to consider the possibility that the transformation of L2 phonetic input

[3] See for instance P. Jusczyk's work in D. Houston, A Seidl, G.Hollich, E. Johnson, & A. Jusczyk (Eds.) (2003). *Jusczyk Lab Final Report*. Retrievable from http://hincapie.psych.purdue.edu/Jusczyk.

[4] Most of this work concerns the relation between alphabetic literacy and phonemic awareness in L1 for young children and non-literate adults.

into L1 output might result not only from perceptual and/or phonological constraints, but also, directly or indirectly, from orthographic influences[5].

1.3. Special features of the Japanese language

In this light, contemporary Japanese is particularly interesting for at least four reasons:

1) Phonologically (Vance, 1987; Itô and Mester, 1995), tautosyllabic consonantal clusters of the /CC/ type are forbidden, as the main syllabic type is /CV/ (or /CyV/)[6], contrary to languages such as French or English. However, we have to distinguish between phonetic and phonological levels, since surface consonantal clusters do exist in Japanese due to vowel devoicing (for instance /musuko/ ('son') pronounced [μ∝σκο] or [μ∝σ∝8κο]).

2) Orthographically (Taylor and Taylor, 1995), the Japanese writing system is composed of one morphophonic system (the kanji) (Matsunaga, 1994), two syllabic (or rather moraic) systems (kana, divided in two scripts, katakana and hiragana) and one alphabetic transliteration system (romaji) based on the Roman alphabet (even though the romaji are not part of the traditional Japanese writing system). The kana system is almost perfectly transparent (one kana = one moraic unit) and the katakana script is usually used to transcribe loanwords.

3) Sociolinguistically, Japanese language has been increasingly incorporating a huge number of loanwords[7], mainly from English, resulting gradually in some form of phonological and orthographic evolution through its stratified phonological lexicon[8] (Fukazawa, Kitahara and Ota, 1998; Itô and Mester, 1999), at least in urban areas like Tokyo. The use of katakana to transcribe these words has been in competition with their original alphabetic spelling, leading to the emergence of new "sounds" like [v][9] and new katakana combinations (Kamyia, 1994), as well as a more frequent use of the roman alphabet.

4) Psycholinguistically (Kess and Miyamoto, 1999), the basic segmentation unit in Japanese is the mora (in Japanese *haku* or *moura*) and not the syllable (in Japanese *onsetsu*), as it appears in the kana writing systems. These kana represent the "natural" segmentation units of speech (Otake, Hatano, Cutler and Mehler, 1993), while the

[5] See also (Chevrot and Malderez, 1999) for a presentation of the so-called "Buben effect" in L1 at the community level, even though it is not directly related to our subject.

[6] All Japanese consonants /C/ have a palatalized equivalent /Cy/ in initial position (Akamatsu, 1997 ; Labrune, 2001a).

[7] Historically, Japanese has always borrowed, especially from Chinese and also Portuguese.

[8] The Japanese lexicon is usually divided into four distinct morphemic strata: Native ("yamato-kotoba"), Sino-Japanese ("kan-go"), Foreign ("gairai-go"), Mimetic / Onomatopoetic ("gitai- / gisei-go"). This stratification appears in the Japanese phonology, morphology and also writing, since, typically, hiragana and kanji are used for the Yamato stratum, kanji for the Sino-Japanese and katakana for the Foreign one. For more details, see for instance (Gelbart and Kawahara, 2006).

[9] Traditionally, [v] was transcribed with the hiragana or katakana characters standing for the Japanese /b/-onset syllable series /ba, bi, bu, be, bo/ : < ば、び、ぶ、べ、ぼ > (hiragana) and < バ、ビ、ブ、ベ、ボ > (katakana). However, to distinguish between /v/ and /b/ in loanwords, the former has been increasingly transcribed with the katakana characters < ヴァ、ヴィ、ヴ、ヴェ、ヴォ > (/va, vi, vu, ve, vo/), a combination of three signs: the katakana character standing for the Japanese vowel /u/ (<ウ>), the voice-marking diacritic sign ("dakuten") (< ゛ >) used to distinguish <カ> (/ka/) and <ガ> (/ga/) for instance, and a half-size katakana vowel character like <ァ> (/a/) for instance, except in the case of /vu/ (<ヴ>).

concept of syllable itself is not as "natural" for Japanese speakers as it can be for French speakers, for instance (Labrune, 2001b).

We selected French as the L2 under scrutiny for two reasons. First, very few French-L2/Japanese-L1 studies exist in the field of interphonology, as opposed to English, which is the main L2-language of reference in this field. This is particularly true in the case of Japanese as an L1, since many linguistic or psycholinguistic studies deal with the divergent negative transfer between Japanese /r/ and English /r, l/, given the phonetic similarities between the rhotic flap and the two liquids. French [®], on the contrary, does not bear such phonetic resemblance, and might therefore be of interest in the case of consonantal clusters of the Obstruant + /r/ type. The second reason is the learners' presumed isolation from French (as compared to English) outside the classroom in Japan, which should help us gain a better insight into the process at work, since any pre-existing sociolinguistic factor (like English loanwords implicit or explicit transcription conventions) would either have a reduced effect or be brought into light by our results.

Given the characteristics of the Japanese language, when Japanese listeners perceive French words containing consonantal clusters, one could expect an "adaptation" of these clusters, in accordance with Trubetzkoy's concept of "phonological filter" (Trubetzkoy, 1939), into less marked structures (Eckman, 1977; Tarone, 1987; Morelli, 1998) of the CV type. In a series of studies (Dupoux, Fushimi, Kakehi and Mehler, 1999; Dupoux, Kakehi, Hirose, Pallier and Mehler, 1999; Dupoux, Pallier, Kakehi and Mehler, 2001; Hirose and Dupoux, 2004), Dupoux and his colleagues actually showed that, in such a case, Japanese listeners "inserted" an epenthetic vowel between the two consonants[10], which the authors interpreted as a pre-lexical influence of their L1 phonotactic knowledge, occurring during the perception of the stimuli. Their position is that loanword adaptations are phonetically minimal transformations that take place in perception (Peperkamp and Dupoux, 2003). These results are incorporated into a speech processing model composed of two levels: a syllabic bank and a segmental bank. It exploits the idea that syllabic segmentation precedes segmental categorization through a matching process of the input with stored syllabic exemplars and explains the influence of syllabic structure on the perception of segments (Dupoux, 2004). This is also congruent with experiments that show that Japanese listeners are sensitive to fine phonetic details (Vendelin and Peperkamp, 2004), since their limited syllable types lead them to pay attention to fine phonetic details.

Even though one cannot assimilate loanword phonology (adaptation of L2 input to L1 grammar) and interphonology (adaptation of L1 grammar to L2 input) (Grijzenhout and Van Rooy, 2001), this epenthetic strategy has been observed in both areas, apparently favored over consonantal deletion (Paradis and Béland, 2002). However, two interrelated points should be mentioned: 1) the role of the orthographic factor in the L2 input perception (Vendelin and Peperkamp, 2005: "Often some standardization has taken place as to how the L2 graphemes are pronounced in the L1. […] We conclude that loanword adaptations made on the basis of the written input are heavily influenced by the L1 speakers knowledge of the L2 orthography"), and 2) sociolinguistic factors regarding the adaptation of non-Japanese syllabic structure (Haunz, 2002: "The methodological issue with Dupoux et al's study is that […] any Japanese speaker is aware of this convention of

[10] See also Akita (2001) and Matthews & Brown (2004) for work in this vein in the case of English as an L2.

adaptation, and this knowledge might play a dominant role in this automatic assumption of a vowel between two consonants by Japanese listeners”). Indeed, due to the massive amount of foreign words in contemporary Japanese, the use of Roman characters has been in competition with the katakana, traditionally used for these transcriptions. Given that many, if not most, foreign words are currently introduced under their orthographic and not their phonetic forms, their phonological adaptation might be influenced by their orthographic presentation. Phonographemic adaptation might therefore be as “standardized” (if not more) as phonetic adaptation, and the role of the auditory (phonetic) perceptual factor in phonological adaptation has to be reexamined in this light.

1.4. Aims of the study

In the case of words containing /CC/ clusters, a visual presentation of these words would leave no doubt as to the presence/absence of a vowel between the two consonants. Therefore, we would expect fewer epenthetic insertions under visual presentation than under auditory presentation. However, given the potentially active “phonographemic conventions” of adaptation of alphabetic foreign words into moraic katakana words and hence simpler phonological structures, one could wonder whether a visual presentation of the word would not activate phonographemic correspondences leading to more epenthetic vowels.

Therefore, the present study had two motivations. First, to examine the influence of Roman orthography on syllabic judgments by Japanese learners of French, by comparing auditory, visual and synchronous audiovisual presentation of non-words containing /CC/ clusters[11]. Second, for each modality, to assess the influence of the nature of the /CC/ clusters on the epenthetic strategy (Broselow, 1983): firstly by comparing /s/ + Stop and Obstruent + Liquid clusters, since this distinction has received particular attention in interphonology literature (the former violate the Sonority Principle) (Durand, 1990; Fleischacker, 2000); secondly by comparing the Obstruent + /l/ and Obstruent + /r/ clusters, as the /l/ *vs*. /r/ distinction is known to be difficult for Japanese speakers (Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Kettermann and Siebert, 2003) under auditory presentation, in English but also in French (it should be noted that the two corresponding graphemes are usually transliterated into a unique kana character).

2. Method

2.1. Participants

62 native Japanese undergraduate university students in Japan, attending different French language classes[12], with a level of French knowledge ranging from complete beginner to intermediate. These were students present in the classes that we chose for our tests, and no selection procedure was involved.

2.2. Stimuli

[11] The “standard” syllable structure in French is of the (C)(C)(C)V(C)(C) type. The most frequent structures are respectively CV, CCV, CVC and V (Wioland, 1991).
[12] Five different classes with 20, 9, 7, 21 and 5 students respectively.

In order to avoid semantics-based top-down effects due to students' prior knowledge or perceptual similarity with Japanese or English cognate words, we created trisyllabic non-word stimuli based on the distribution of the following French tautosyllabic consonantal clusters given by Dell (1995) in initial, medial and final word-position (for instance: /trosema, sematro, semagotr/):

OBLI: /pl, pr, bl, br, fl, fr, vl, vr, tr, dr, kl, kr, gl, gr/.
/s/ + Stop: /sp, st, sk/.
/s/ + Nonstop: /sl, sm, sn, sv, sf/ (but only /sm/ in final position).

For the auditory stimuli, fillers were created by inserting an epenthetic vowel (mainly /u/, and /o/ after /t/ and /d/) between the two consonants of each non-word (for instance /plokama/ and /pulokama/), and the two lists (trisyllabic target stimuli and quadrisyllabic and pentasyllabic (in word-final position) filler stimuli) were randomized (total = 124 stimuli). For the visual stimuli, only five fillers were inserted in the original list, evenly distributed (total = 67 stimuli). As for the audiovisual presentation, half of the stimuli were congruent (same length for corresponding audio and visual stimuli) and half were not (one longer than the other because of an epenthetic vowel and vice-versa) (total = 186 stimuli). The audio stimuli were recorded in a soundproof studio by one of us on a MD recorder and then copied onto an audio tape, as is frequently done by teachers who record material for their classes. They were played in class with the tape player usually used by the teacher, with two loudspeakers facing the group of students. The visual stimuli were printed in black capital letters (Normal, Times New Roman, 65) on a white sheet of paper and projected on a screen with the OHP usually used in the classroom. For each condition (auditory, visual and audiovisual), the order of presentation of the stimuli was different.

2.3. Procedure

The tests were presented to the students as learning activities in their usual learning environment (space, time, conditions and teacher). The students were provided with answer sheets and initial explanations with examples in Japanese. The answer sheets displayed a table with a line for each stimulus. For each line, there were three columns with three, four or five ellipses (OOO / OOOO/ OOOO). The students were asked to circle the number of ellipses corresponding to the number of syllables (in Japanese *onsetsu*) in the words they heard or read[13]. They were asked to answer "without thinking" as there was no "good" or "wrong" answer. The experiment (group testing) was run during a 1h30 course for each class, first in auditory, then in audiovisual and finally in visual condition. All stimuli (auditory and visual) were separated by a beeping sound recorded on a tape and occurring every 7 seconds, which implies that the visual stimuli

[13] For a discussion of the status of mora and syllable in Japanese, see Otake, Hatano, Cutler & Mehler (1993), Kubozono (1989) and Labrune (2001b). Whether the subjects would count the latter or the former somehow was part of the issue at stake. However, it could only be really relevant in the case of the nasal mora /N/ and hence for the cluster /sn/ in our protocol. Yet, even if the mora /N/ was counted as one unit in our stimuli (however implausible it seems given the stimuli : snotipa / sunotipa & tipasno / tipasuno), the protocol would still be valid and our measures reliable.

were accessible a longer time than the auditory ones, since each sheet of paper was manually removed by the teacher. For the audiovisual condition, students were informed that there might be discrepancies between the two modalities, and an extra column was provided to signal the perception of such discrepancies. Thus, the task in audiovisual condition was more complex, since students had to count syllables but also to notice possible discrepancies between the two sources. This was designed to ensure that students would not ignore one of the two sources of input. Likewise, in the visual condition, the five filler stimuli were designed to ensure that students would not adopt a systematic answer strategy and would actually count the number of syllables for each written stimulus.

3. Results

The results presented here are based on the number of errors that reflected epenthesis[14]: when the number of ellipses chosen on the sheet was superior to the number of syllables in the stimulus, it was counted as one case of epenthesis, even in the case of coda clusters which theoretically might lead to a double epenthesis.

3.1. Modality effect

For the same stimuli, differing only in modality (auditory (A), audiovisual (AV) and visual (V)), order of presentation and filler types, we obtained the following descriptive results, when all three types of /CC/ clusters were taken into account:

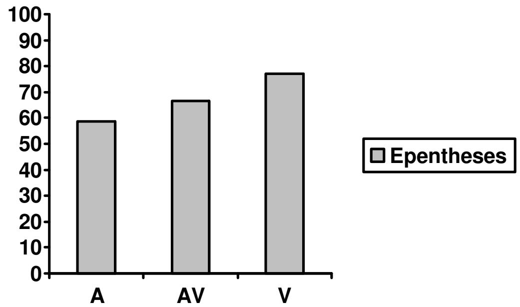
Table 1
Global mean epenthesis rates in A, AV and V condition

	A condition	**AV condition**	**V condition**
Mean epenthesis rate	58.5%	66.5%	77.03%
Standard deviation	25.63%	27.8%	30.96%

As appears in Fig. 1, the rate of epenthesis increases with the visual modality, whether in AV or in V only condition.

Fig. 1
Global epenthesis rates in A, AV and V condition

[14] We decided to use the words “epenthesis” and “deletion” even though the procedure did not involve actual production.



This trend has been found in all three positions (word-initial, medial and final). A particular increase has been observed in word-final position (approximately + 10% between initial and final positions in both AV and V conditions, whereas it was more stable in the A condition). This inter-modality difference might be explained by the visual salience of codas in AV and V conditions, otherwise less salient in A condition. The visual presentation of the stimuli might force subjects to actually perceive complex codas (and therefore to "repair" them (LaCharité and Paradis, 1993) with epenthetic vowels), contrary to auditory presentation which, in spite of French final-syllable stress, might allow more perceptual deletion of one of the two consonants (possibly the less sonorous)[15]. The positional variation in AV and V conditions might be accounted for by the potentially double epenthetic insertion in word-final complex codas, as opposed to single epenthesis in initial and medial-word position complex onsets, which increases the potential for epenthesis in coda position.

3.2. Cluster nature effect

The distinction between /s/ + Stop, /s/ + Nonstop, and OBLI leads to the following results:

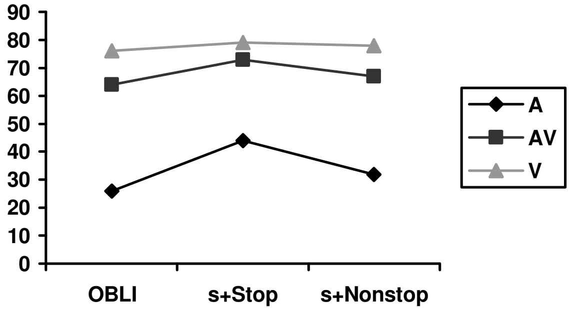
Table 2
Mean epenthesis rates in A, AV and V condition for OBLI, /s/ + Stop and /s/ + Nonstop clusters

/CC/ clusters	OBLI	/s/ + Stop	/s/ + Nonstop
A condition	26%	44%	32%
AV condition	64%	73%	67%
V condition	76%	79%	78%

When we consider Fig. 2, it appears that, as in Fig.1, the epenthesis rate increases with the visual modality for the three types of clusters.

Fig. 2
Mean epenthesis rates in A, AV and V condition for OBLI, /s/ + Stop and /s/ + Nonstop clusters

[15] On sonority and consonant deletion see Tropf (1987). See also De Cara (1998) on sonority profile.



However, the rate difference between the three types of clusters seems to decrease with the visual modality. This could indicate that the phonetic properties of the stimuli (which are lost in the visual condition) might play a role in syllabic perception in auditory condition, while visual stimuli activate phonological representations that would match the Japanese syllabic structure and be less faithful to the input. When we checked the audiovisual results to ensure that the subjects were not systematically biased towards one modality over the other, through the analysis of the divergent stimuli perception results, we found no clear superiority of one over the other, even though the visual modality might have been slightly predominant (which could easily be attributed to the fact that the visual stimuli were accessible for a longer time than the auditory ones). Even if that was the case, it would be congruent with our general results.

3.3. OBLI effect

While the figures presented in Table 2 showed more epentheses in /s/ + Stop than in both the other two conditions (OBLI and /s/ + non-stop), the results of the /l/ *vs*. /r/ distinction in OBLI clusters are shown in Table 3:

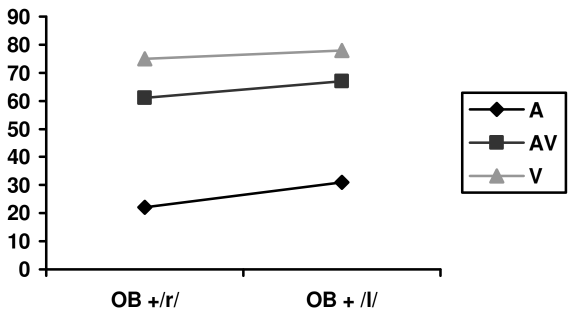
Table 3
Mean epenthesis rates in A, AV and V condition for OB + /r/ and OB + /l/ clusters

/CC/ type	**A condition**	**AV condition**	**V condition**
OB+ /r/	22%	61%	75%
OB+/l/	31%	67%	78%

These figures confirm the previous trend and also indicate a higher rate of epenthesis for OB + /l/ over OB + /r/ clusters in all three conditions[16].

Fig. 3
Mean epenthesis rates in A, AV and V condition for OB + /r/ and OB + /l/ clusters

[16] As one of the two anonymous reviewers suggested, it is worth mentioning the studies of Riney (1990) and Weinberger (1987) on variable epenthesis, because written exposure is hardly mentioned as source of variation.



As in Fig. 2, the rate difference between the two types in Fig. 3 decreases with the visual modality.

4. Discussion

First of all, the results in the auditory condition fit into what has been previously reported in phonological literature, namely the special status of the /s/ + Stop clusters, which violate the Sonority Principle. In the case of Japanese, vowel devoicing has to be considered to explain these results (Vance, 1987), since the vowel /u/ is often devoiced between /k, s, t, p/, especially when /u/ is preceded by /s/. Therefore, phonetic clusters such as [sp, st, sk], which exist in Japanese, would correspond to /sup, sut, suk/ phonological sequences for Japanese speakers, while OBLI and /s/ + non-stop phonetic clusters would not (since such phonetic clusters do not exist in Japanese). In a perceptually-driven metaphonological task such as ours, one might therefore expect more epentheses (as we defined them, and not as actual vocal epenthetic production) with /s/ + Stop clusters than with the two others, since listening to [s] + [Stop] clusters would *automatically* (from a psycholinguistic viewpoint) activate /su/ + Stop phonological sequences in Japanese, whereas it is not the case for the two others, for which *controlled* (possibly more input-faithful) perceptual process would be less influenced by such phonetic-phonological mapping correspondences in Japanese.

For the OBLI clusters, the distinction between /l/ and /r/ revealed that OB+/l/ clusters triggered more epentheses than OB + /r/. Following the same type of argument as the preceding one[17], this could be explained by the fact that [l] is an allophonic realization of the Japanese /r/ (Akamatsu, 1997), leading to the identification of a phonetic stimulus like [kla] as the phonological sequence /kura/ (with vowel devoicing, even though it does not compare with [s] + Stop clusters, which do exist phonetically in Japanese). On the contrary, stimuli containing French [®] could lead either to consonantal deletion (/ka/) or

[17] These possible explanations have to be understood within the frame of our experimental protocol. One could consider that since phonetic clusters [sC_1] do exist in Japanese (in oral production) whereas [sC_2] clusters do not, one might expect more epentheses for [sC_2], and then for /sC_2/. That would fit an *oral production* task. However, in our syllabic counting task, the issue is one of mapping between phonetic / graphic perceptual input and phonological / phonographemic representations. Given that [sC_1] clusters usually corresponds to /sVC_1/ in Japanese (automatic process), one might expect more "phonological epentheses" for /sC_1/ than for /sC_2/, since [sC_2] stimuli are totally unknown to the subjects and therefore in need of more controlled processing (which in turn might lead to increased faithfulness to the input).

phonetic misperception (for instance [kHa]), due to the phonetic nature of the French /r/[18]. These results could be associated with the idea that "new sounds" ([®]) are actually less difficult to acquire than known phonetic categories ([l]) (Flege, 1990), due to perceived phonetic dissimilarities (Aoyama, Flege, Guion, Akahane-Yamada and Yamada, 2004) and L1-transfer effects (Escudero and Boersma, 2004).

Prior to the experiment, we expected a smaller rate of epenthesis in visual condition under the assumption that the perception of two letters side-by-side would prevent subjects from inserting an epenthetic vowel in the task of syllable counting. On the contrary, the visual condition turned out to trigger the highest rate of epenthesis, and the figure (77%) made us wonder about any potential methodological flaw. After a careful examination of the results for each subject, we found no such flaw. Instead, in the visual condition, the answers were quite consistent: some subjects performed very well (*almost* no epenthesis) while the others were *almost* systematically wrong in their syllabic counting[19]. Moreover, when we compare our results with those of Dupoux, Fushimi, Kakehi and Mehler (1999) in a task of transcription in which subjects inserted an epenthetic <u> for 83% and 74% of the stimuli, our data seem to reach the same proportions.

---

[18] Besides, when splitting out the voiceless and voiced OBLI clusters, we found the following results in A condition (almost no difference in V condition), which are congruent with our argument:

1) Voiceless OB + /l/ clusters trigger more epentheses than voiced OB + /l/ clusters. That corresponds to the devoicing of /u/ in voiceless environment, which leads to more epentheses in the syllabic counting task, as explained above.

2) Voiced OB + /r/ clusters trigger more (or as many) epentheses than (as) voiceless OB + /r/ clusters. This could be accounted for by the fact that, given the non-existence of [®] in Japanese, its voiceless allophonic realization when preceded by a voiceless OB (usually a voiceless uvular fricative [Ξ]) might be more frequently (phonetically) misperceived than its voiced counterpart, and therefore trigger fewer epentheses. This argument is reinforced by the fact that the highest difference rate has been found between the clusters /tr/ and /dr/ (and not /pr, br/ or /kr, gr/), while the epenthetic vowel usually found between the two coronals and /r/ (which we used in our protocol) is /o/ (and not /u/, as is the case for the four other clusters), which is generally not devoiced in Japanese. Taking that into account, the fact that /tr/ triggers fewer epentheses than /dr/ supports the idea of more frequent misperception of [Ξ] as a segmental unit and fits into our argument.

3) In the case of /f/ and /v/, the voiced clusters (/vl/ and /vr/) always trigger more epentheses than their voiceless counterparts. Beside what has been mentioned above, the difference between /f/ and /v/ has to be linked with the special status of /v/ in Japanese: while neither of them exist in the traditional Japanese segmental inventory (/f/ would be adapted as /h/, phonetically realized as a bilabial fricative [÷]), /v/ can be considered as an emerging "new consonant" in Japanese (Labrune, 2001a, p. 94), imported through loanwords, and the Japanese intervocalic allophonic realization of /b/ as a bilabial [B] has always been very similar phonetically to the French [v]. Therefore, given the relative existence of the segment /v/ (or at least its phonetic realization) in contemporary Japanese, it seems that the [v] clusters would be more easily perceived as a bisegmental unit than the [f] ones, and thus trigger more epentheses.

[19] Even though these results might cast doubts on the actual performance of the subjects in that task, the fact that the two groups were not *completely* either wrong or right indicate that they actually did what was expected from them, i.e. approximately (since the instruction was given in Japanese) "Read the letters shown on the screen, divide the word in syllables ("onsetsu") and circle the box which has the corresponding number of ellipses". Therefore, the distinction between "spoken realization" and "visible syllable" (which would have to be defined) one might be tempted to think of is not relevant here, since the task was precisely designed to test the metaphonological awareness of the subjects, through the activation of phonological representations, as it was in the A condition.

In the audiovisual condition, the double input, the instructions and the incongruent stimuli might have triggered attentional "trade-off'. However, our protocol was designed to ensure subjects would not be biased towards one modality over the other. Besides, in spite of the higher complexity of the task in the AV condition, the epenthesis rate remains lower than the one in the V condition. Given that the epenthesis rate increases according to an A < AV < V scale for all clusters, we can consider that, in the audiovisual condition, subjects might have alternatively focused on one or the other modality, or assessed each source of input (Massaro, Cohen and Smeele, 1996) to reach a phonological judgment (in the AV condition, a small percentage of answer was congruent neither with the auditory nor with the visual stimuli). Therefore, we can focus on the comparison between A and V conditions.

When we compare the results we obtained in auditory and visual conditions, we might be inclined to think that the phonetic properties of the auditory stimuli influenced the phonological judgment of the subjects (which reveals the phonetic sensitivity of the listeners (Vendelin and Peperkamp, 2004)), while visual stimuli activate phonological representations more attuned to the phonological and orthographic structure of the L1. From a general viewpoint, six explanatory perspectives can be considered and combined.

First, from a mnesic viewpoint, according to Baddeley's Working Memory model (Baddeley, 1986), in which the phonological loop component[20] plays a role in lexical and phonological learning (Papagno, Valentine and Baddeley, 1991; Baddeley, Gathercole and Papagno, 1998), it appears that for non-expert readers, visual information has to be recoded with the available phonological codes in order to be stored, whereas auditory information could be directly maintained and refreshed in the phonological loop[21]. Therefore, orthographic input might be more influenced by the moraic structure of the Japanese phonological and orthographic system than phonetic input.

Second, from a metaphonological viewpoint, if we consider that the syllable counting task strongly involves the subjects' phonological awareness, we have to remember that this awareness is partially shaped by the writing system of our literate subjects and that written representations are fundamentally speech segmentation tools. In such a task, mental orthographic representations might therefore be activated by both auditory and visual stimuli. Given the level of our subjects in French, it is highly probable that these representations will be kana-based (and not alphabet-based), especially when we consider that these kana are fundamentally the symbolic moraic units of Japanese language segmentation. However, even if the first representations (audio-activated) are partially determined by the L1 phonographemic system, the others (visually-activated) are fully determined by these phonographemic correspondences since their conversion into a mental phonological representation is crucially dependent on them. Our results could support the idea that the phonographemic correspondences activated by the visual stimuli lead to a phonological representation shaped by L1 phonological and

[20] In his 1986 book, Baddeley indicates (p. 75) : « I shall use the term "phonological" in a purely neutral sense as meaning speech-based. The term "articulatory" will be used when the coding is assumed to be based on the subject's speech production system, while the term "acoustic" will be used when the coding is assumed to be based on the subject's speech perception processes ».

[21] Which does not mean there is no phonological analysis nor abstract phonological representation. For more details see (Baddeley, 2003, p. 193) and (Repovs & Baddeley, 2006, p. 7: "while speech input enters the phonological store automatically, information from other modalities enters the phonological store only through recoding into phonological form, a process performed by articulatory rehearsal").

phonographemic constraints, while auditory stimuli would be less constrained by the phonographemic system. However, to be fully developed, this argument has to be combined to the next perspective, which deals with the alphabet/kana conversion codes.

Third, from a sociophonological viewpoint, we mentioned earlier that in contemporary Japanese, especially among younger subjects like our university students, the traditional plurality of the Japanese writing system seems to be expanding to include the Roman alphabet, due to the massive use of loanwords, essentially from English (Kamiya, 1994; Kess and Miyamoto, 1999; Taylor and Taylor, 1995). These words are mainly used in their written form in the mass media, and they now tend to be retained in their original orthographic form, whereas they used to be either transcribed in katakana or in an adapted romaji form. Both forms (katakana and alphabet) are commonly associated on signs and magazines, so that, gradually, phonographemic correspondences between alphabet, katakana and Japanese-adapted phonological forms seem to become standardized, at least among urban Japanese. An orthographic <CC> type stimulus would therefore automatically activate a phonological /CvC/ form for such Japanese speakers, as if a new orthographic system was introduced without introducing new phonological units [22]. Therefore, this sociolinguistic factor sheds a new light on the preceding metaphonological argument: while visual stimuli would activate *automatic* conversion codes, auditory stimuli would necessitate a more *controlled* process, enabling non-standard segmentation of the input.

Fourth, from a phonetico-phonological viewpoint, we have to consider the distinction between *phonetic* and *phonological* syllables (Blevins, 1995), given the existence in Japanese of the vowel devoicing phenomenon. Indeed, in auditory condition, subjects might be able to count "phonetic syllables", whereas visual stimuli would activate phonological representations (and syllables) in which the "devoiced" vowel would be counted on the basis of its temporal slot, and therefore lead to more epentheses. This argument has to be associated to the next two perspectives.

Fifth, from a perceptual viewpoint, we could distinguish auditory and visual stimuli on the basis of their inherent characteristics. Contrary to visual stimuli, auditory stimuli are essentially elusive and their perception can be influenced by various acoustic and prosodic parameters as well as surrounding "noise", especially in the real-class group-testing setting of our experiment. According to this, "misperception" of /CC/ clusters in auditory condition, leading to consonantal deletion and therefore to a lesser rate of epenthesis, would be more frequent than in visual condition[23].

Finally, from an attentional viewpoint, we mentioned earlier that, given the characteristics of the auditory stimuli, subjects might have to pay more attention to the auditory (controlled process) than to the visual (automatic process) stimuli. However, despite this allocation of attentional resources, we also indicated that "misperception" of the /CC/ clusters, for instance through consonantal deletion, might be more frequent in the auditory than in the visual condition. Thus, the visual condition would lead to a higher degree of preservation of the input (no deletion due to perceptual salience) as is shown in the case of the coda mentioned in the results, while at the same time triggering more epentheses (due to automatic process). It should also be mentioned that the visual

[22] Or constraint re-ranking in an OT framework for instance.
[23] As one of our anonymous reviewers suggested, it would have been useful to include in the experiment, beside syllabic segmentation, a count of the number of segments/sounds respondents could hear.

condition was the last one in the series of our tests, which could lead to a diminution of attentional control by the subjects. On the contrary, in the auditory condition, consonantal deletions could be more frequent (due to perceptual noise), while epentheses could be less systematic (due to controlled process). It could then be considered that in the processing of /CC/ clusters stimuli, vowel epentheses should be predominant in the visual condition, while consonantal deletions should not be neglected in the auditory condition.

5. Conclusion

Combined together, these six perspectives seem to account for our results. These eventually point to the link between orthographic representations and "deep" phonological representations, as opposed to phonetic surface form. They suggest that orthography should be more systematically taken into account in speech perception research, in loanword phonology and interphonology studies (Detey, 2005; Detey, Durand and Nespoulous, 2005a; Detey, Durand and Nespoulous, 2005b), either in the experimental design or in the interpretation of the results. This point was already mentioned by Damian and Bowers about the priming effect technique (2003): "The numerous studies that have employed this technique have not considered the relevance of spelling, and accordingly, effects attributed to phonology may have another source". Thus, in spite of the obvious methodological weaknesses of our study, our results seem to be convergent with the relatively recent focus on the role of orthography in speech processing studies. Our tests should therefore be replicated in a more controlled environment to further our preliminary approach of the issue.

References

Akamatsu, T., 1997. Japanese Phonetics: theory and practice. Lincom Europa, Munich.

Akita, M., 2001. The Phonological Acquisition of Japanese Learners of English : a longitudinal Study of Perception and Production. Unpublished Ph.D. dissertation, University of Durham.

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., Yamada, T., 2004. Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /r/. Journal of Phonetics 32, 233-250.

Archangeli, D., 1997. Optimality Theory : an introduction to linguistics in the 1990s. In: Archangeli, D., Langendoen, D. T. (Eds.), Optimality Theory : an overview. Blackwell, Maden, pp. 1-33.

Baddeley, A. D., 1986. Working Memory. Oxford University Press, Oxford.

Baddeley, A. D., 2003. Working memory and language: an overview. Journal of Communication Disorders 36, 189-208.

Baddeley, A. D., Gathercole, S., Papagno, C., 1998. The phonological loop as a language learning device. Psychological Review 105 (1), 158-173.

Bassetti, B., 2006. Orthographic conventions and phonological representations in learners of Chinese as a Foreign Language. Written Language and Literacy 9 (1), 95-114.

Bentin, S., 1992. Phonological awareness, reading, and reading acquisition: a survey and appraisal of current knowledge. In: Frost, R., Katz, L. (Eds.), Orthography, Phonology, Morphology and Meaning. Elsevier Science, Amsterdam, pp. 193-210.

Blevins, J., 1995. The syllable in phonological theory. In: Goldsmith, J. (Ed.), The Handbook of Phonological Theory. Blackwell, Oxford, pp. 206-243.

Broselow, E., 1983. Nonobvious transfer: on predicting epenthesis errors. In Gass, S., Selinker, K. (Eds.), Language Transfer in Language Learning. Newbury House, Rowley, pp. 269-280.

Byung-jin, L., 2004. Effects of L1 phonology and orthography on L2 perception: the use of Korean Hangul in perception of English voicing contrasts. Paper presented at 9th Conference on Laboratory Phonology, Jun. 24-26, University of Illinois, Urbana, Illinois.

Chéreau, C., Hallé, P. A., Segui, J., 1999. Interference between surface form and abstract representation in spoken word perception. Eurospeech'99, 1667-1670.

Chevrot, J.-P., Malderez, I., 1999. L'effet Buben : de la linguistique diachronique à l'approche cognitive (et retour). Langue Française 124, 104-125.

Cross, N., 2002. Bilinguals' and Second Language Learners' Knowledge of Japanese Syllable Structure. Unpublished Ph.D. dissertation. University of Durham.

Cutler, A., Mehler, J., Norris, D., Segui, J., 1986. The syllable's differing role in the segmentation of French and English. Journal of Memory and Language 25, 385-400.

Damian, M. F., Bowers, J. S. 2003. Effect of orthography on speech production in a form-preparation paradigm. Journal of Memory and Language 49, 119-132.

De Cara, B., 1998. Profil de sonorité et traitement de la syllabe chez l'apprenti-lecteur. Ph.D. dissertation, Université de Bourgogne, ANRT.

Dell, F., 1995. Consonant clusters and phonological syllables in French. Lingua 95, 5-26.

Detey, S., 2005. Interphonologie et représentations orthographiques. Du rôle de l'écrit dans l'enseignement/apprentissage du français oral chez des étudiants japonais. Unpublished Ph.D. dissertation, Université de Toulouse-Le Mirail, 1309 p.

Detey, S., Durand, J., Nespoulous, J.-L., 2005a. Apprentissage de la phonologie du français par des Japonais et multimodalité: une didactique nécessairement cognitive. In Billières, M., Gaillard, P., Spanghero-Gaillard, N. (Eds.), Actes du Premier Colloque International de Didactique Cognitive, (Français Langue Etrangère / Langue Seconde / Langue Maternelle) [CD-ROM]. Université de Toulouse-le Mirail, Toulouse, pp. 144-149.

Detey, S., Durand, J., Nespoulous, J.-L., 2005b. Interphonologie et représentations orthographiques. Le cas des catégories /b/ et /v/ chez des apprenants japonais de français langue étrangère. Revue PArole 37/38/39 (supplément), 139-186.

Dijkstra, T., Frauenfelder, U. H., Schreuder, R., 1993. Bidirectional grapheme-phoneme activation in a bimodal detection task. Journal of Experimental Psychology : human perception and performance 19 (5), 931-950.

Dupoux, E., 2004. The acquisition of discrete segmental categories: data and model. Paper presented at 18th International Congress of Acoustics, Apr. 4-9, Kyoto.

Dupoux, E., Fushimi, T., Kakehi, K., Mehler, J., 1999. Prelexical locus of an illusory vowel effect in Japanese. Eurospeech'99, 1675-1678.

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.

Dupoux, E., Pallier, C., Kakehi, K., Mehler, J., 2001. New evidence for pre-lexical phonological processing in word recognition. Language and Cognitive Processes 16 (5-6), 491-505.

Durand, J., 1990. Generative and Non-Linear Phonology. Longman, London.

Durand, J., 2000. Oral, écrit et faculté de langage. In: Guillot, M.-N., Kenning, M.-M. (Eds.), Changing Landscapes in Language and Language Pedagogy. AFLS/CILT, London, pp. 40-72.

Eckman, F. R., 1977. Markedness and the contrastive analysis hypothesis. Language Learning 31, 195-216.

Eckman, F. R., 2004. From phonemic differences to constraint rankings: research on second language phonology. Studies in Second Language Acquisition 26 (4), 513-549.

Escudero, P., 2004. Revisiting the role of perception in segmental phonology. In: Kerkhoff, K., Lange (de) J., Sadeh Leicht, O. (Eds.), UiL OTS Yearbook 2004, Utrecht Institute of Linguistics, Utrecht, 2004.

Escudero, P., Boersma, P., 2004. Bridging the gap between L2 speech perception research and phonological theory. Studies in Second Language Acquisition 26 (4) 551-585.

Flege, J. E., 1990. English vowel production by Dutch talkers: more evidence for the “similar” vs. “new” distinction. In: James, A., Leather, J. (Eds.), New sounds 90: Proceedings of the Amsterdam Symposium on Acquisition of Second-Language Speech, University of Amsterdam, Amsterdam, pp. 255-293.

Fleischacker, H., 2000. Cluster-dependent epenthesis asymmetries, Unpublished ms, UCLA.

Fukazawa, H., Kitahara, M., Ota, M., 1998. Lexical stratification and Ranking invariance in constraint-based grammars. Proceedings of the CLS 34, Rutgers Optimality Archives 267-0698.

Gelbart, B., Kawahara, S. 2006. Psychological reality of sublexica in Japanese. Paper presented at the 4th Formal Approaches to Japanese Linguistics Conference, University of Osaka.

Grainger, J., Nguyen Van Kang, M., Segui, J., 2001. Cross-modal repetition priming of heterographic homophones. Memory and Cognition 29 (1), 53-61.

Grijzenhout, J., Van Rooy, B., 2001. Second language phonology: acquisition through gradual constraint demotion. Working Paper. Theorie des Lexicons. University of Düsseldorf.

Haunz, C., 2002. The role of speech perception in the adaptation of loanwords. Paper presented at Post-Graduate Conference (PGC) 2002, University of Edinburgh.

Him Cheung, H.-C., Chen, H.-C., 2004. Early orthographic experience modifies both phonological awareness and on-line speech processing. Language and Cognitive Processes 19 (1), 1-28.

Hirose, Y., Dupoux, E., 2004. Ignoring the signal: evidence from vowel epenthesis. Poster paper presented at 10th Annual Conference on the Architectures and Mechanisms for Language Processing (AMLaP), Sept. 16-18, Aix-en-Provence, France.

Itô, J., Mester, R. A., 1995. Japanese phonology. In: Goldsmith, J. (Ed.), The Handbook of Phonological Theory. Blackwell, Oxford, pp. 817-835.

Itô, J., Mester, R. A., 1999. The phonological lexicon. In: Tsujimura, N. (Ed.), The Handbook of Japanese Linguistics. Blackwell, Oxford, pp. 62-100.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A. Siebert, C., 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. Cognition 87, B47-B57.

Jacobs, H., Gussenhoven, C., 2000. Loan phonology: perception, salience, the lexicon and OT. In: Dekkers, J., Leeuw, F. van der, Weijer, J. van de (Dirs.), Optimality Theory. Oxford University Press, Oxford, pp. 193-209.

Kamiya, T., 1994. Tuttle New Dictionnary of Loanwords in Japanese. Charles E. Tuttle, Rutland (Vermont).

Kess, J. F., Miyamoto, T., 1999. The Japanese Mental Lexicon: psycholinguistic studies of kana and kanji processing. John Benjamins, Amsterdam.

Kubozono, H., 1989. The mora and the syllable structure in Japanese : evidence from speech errors. Language and Speech 32, 249-278.

Labrune, L., 2001a. Phonologie du japonais. In : Blin, R., Tamba, I. (Eds.), Faits de Langues « coréen-japonais ». Ophrys, Paris, pp. 89-110.

Labrune, L., 2001b. Structure de la syllabe japonaise. In : Blin, R., Tamba, I. (Eds.), Faits de Langues « coréen-japonais ». Ophrys, Paris, pp. 112-122.

LaCharité, D., Paradis, C., 1993. The emergence of constraints in generative phonology and a comparison of three current constraint-based models. The Canadian Journal of Linguistics 38 (2), 127-150.

Massaro, D. W., Cohen, M. M., Smeele, P. M. T., 1996. Perception of asynchronous and conflicting visual and auditory speech. Journal of the Acoustical Society of America 100 (3), 1777-1786.

Matsunaga, S., 1994. The Linguistic and Psycholinguistic Nature of Kanji: do kanji represent and trigger only meanings? Ph.D. dissertation, University of Michigan, Ann Arbor.

Matthews, J. Brown, C. 2004. When intake exceeds input: language specific perceptual illusions induced by L1 prosodic constraints. International Journal of Bilingualism 8 (1), 5-27.

Mehler, J., Dommergues, J. Y., Frauenfelder, U., 1981. The syllable's role in speech segmentation. Journal of Verbal Learning and Verbal Behavior 20, 298-305.

Morais, J., Cary, J., Alegria, J., Bertelson, P., 1979. Does awareness of speech as a sequence of phones arise spontaneously ? Cognition 7, 323-331.

Morais, J., Content, A., Cary, L., Mehler, J., Segui, J., 1989. Syllabic segmentation and literacy. Language and Cognitive Processes 4, 57-67.

Morelli, F., 1998. Markedness relations and implicational universals in the typology of onset obstruent clusters. Proceedings of North Eastern Linguistics Society 28 (2), pp. 107-120.

Otake, T., Hatano, G., Cutler, A., Mehler, J., 1993. Mora or Syllable ? Speech segmentation in Japanese? Journal of Memory and Language 32, 258-278.

Papagno, C., Valentine, T., Baddeley, A., 1991. Phonological short-term memory and foreign-language vocabulary learning. Journal of Memory and Language 30, 331-47.

Paradis, C., Béland, R., 2002. Syllabic constraints and constraint conflicts in loanword adaptations, aphasic speech, and children's errors. In: Durand, J., Laks, B. (Eds.), Phonetics, Phonology and Cognition. Oxford University Press, Oxford. pp. 191-225.

Pennington, M. C., 1996. Phonology in English Language Teaching. London, Longman.
Peperkamp, S., Dupoux, E., 2003. Reinterpreting loanword adaptations: the role of perception. In: Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, Causal productions.
Ramus, F., Dupoux, E., Mehler, J., 2003. The psychological reality of rhythm classes: perceptual studies. In: Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, Causal productions.
Repovs, G., Baddeley, A., 2006. The multi-component model of working memory: explorations in experimental cognitive psychology. Neuroscience 139, 5-21.
Riney, T., 1990. Age and open syllable production in interlanguage phonology. In: Burmeister, H., Rouends, P. (Eds.), Variability in Second Language Acquisition. Proceedings of the 10th SLRF Meeting, Vol. 2. University of Oregon, Eugene, pp. 655-666.
Rose, Y., Demuth, K., 2006. Vowel epenthesis in loanword adaptation: phonological and phonetic considerations. Lingua 116, 1112-1139.
Segui, J., 1997. La perception du langage parlé: données et théories. In: Lambert, J., Nespoulous, J.-L. (Dirs.), Perception auditive et compréhension du langage: état initial, état stable et pathologie. Solal Editeurs, Marseille, pp. 15-22.
Segui, J., Ferrand, L., 2002. The role of syllable in speech perception and production. In: Durand, J., Laks, B. (Eds.), Phonetics, Phonology and Cognition. Oxford University Press, Oxford. pp. 151-167.
Shinohara, S., 1997. Analyse phonologique de l'adaptation japonaise de mots étrangers. Ph.D. dissertation, Université de la Sorbonne-Nouvelle Paris-III.
Smith, J. L., 2005. Loan phonology is not all perception: evidence from Japanese loan doublets. In: Vance, T. J., Kimberly, A. J. (Eds.), Japanese/Korean Linguistics 14, CSLI, Palo Alto, pp. 63-74.
Tarone, E., 1987. The phonology of interlanguage. In: Ioup, G., Weinberger, S. H. (Eds.), Interlanguage Phonology: the acquisition of a second language sound system. Newbury House, Cambridge, pp. 70-85.
Taylor, I., Taylor, M. M., 1995. Writing and Literacy in Chinese, Korean and Japanese. John Benjamins, Amsterdam.
Tropf, H., 1987. Sonority as a variability factor in second language phonology. In: James, A. and Leather, J. (Eds.), Sound Patterns in Second Language Phonology. Foris Publications, Dordrecht, pp 173-191.
Trubetzkoy, N. S., 1939. Grundzüge der Phonologie. Travaux du cercle linguistique de Prague 7, Prague. Translated in French by J. Cantineau, *Principes de phonologie*, Paris, Klincksieck, 1st edition, 1949.
Uffmann, C., 2003. Markedness vs. Transfer in SLA: a unified account. Paper presented at XVIth International Conference on Foreign/Second Language Acquisition. May 22, Szczyrk.
Vance, T. J., 1987. An Introduction to Japanese Phonology. State University of New York Press, New York.
Vendelin, I., Peperkamp, S., 2004. Evidence for phonetic adaptation of loanwords: an experimental study. In: Actes des Journées d'Etudes Linguistiques 2004 (JEL2004), Université de Nantes, pp. 129-131.

Vendelin, I., Peperkamp, S., 2005. The influence of orthography on loanword adaptations. Paper presented at 7èmes Rencontres Internationales du Réseau Français de Phonologie, Université d'Aix-en-Provence, Jun. 2-4.

Ventura, P., Kolinsky, R., Brito-Mendes, C., Morais, J., 2001. Mental representations of the syllable internal structure are influenced by orthography. Language and Cognitive Processes 16 (4), 393-418.

Weinberger, S. 1987. The influence of linguistic context on syllable simplification. In : Ioup, G, Weinberger, S. (Eds.), Interlanguage Phonology. Newbury House, Cambridge, pp. 401-417.

Wioland, F., 1991. Prononcer les mots du français. Hachette, Paris.

Young-Scholten, M., 2002. Orthographic input in L2 phonological development. In: Burmeister, P., Piske, T., Rohde, A. (Eds.). An Integrated View of Language development – Papers in Honour of Henning Wode. WVT Wissenschaftlicher Verlag, Trier, pp. 263-279.

Ziegler, J. C., Muneaux, M., Grainger, J., 2003. Neighborhood effects in auditory word recognition: phonological competition and orthographic facilitation. Journal of Memory and Language 48, 779-793.