The role of syllable structure in lexical segmentation: Helping listeners avoid mondegreens

Content, Alain; Dumay, Nicolas; Frauenfelder, Ulrich Hans

THE ROLE OF SYLLABLE STRUCTURE IN LEXICAL SEGMENTATION: HELPING LISTENERS AVOID MONDEGREENS

Alain Content, Nicolas Dumay
Laboratoire de Psychologie Expérimentale, Université libre de Bruxelles, Belgique
and **Uli H. Frauenfelder**
Laboratoire de Psycholinguistique Expérimentale, Université de Genève, Suisse

ABSTRACT

One challenge for theories of word recognition is to determine how the listener recovers the intended lexical segmentation in continuous speech. We argue that syllable structure provides one source of constraint on lexical segmentation and more precisely, that syllable onsets constitute potential alignment points for the mapping process. We present an overview of several studies using explicit syllable segmentation tasks, word spotting and crossmodal priming, which support the hypothesis.

1. INTRODUCTION

"Surely you've heard of mondegreens — mishearings of common words and phrases that make their way into speech. It was a British writer in the mid-1950s who admitted that she misheard

They have slain the Earl of Moray
And laid him on the green

as *And Lady Mondegreen*. She could never figure out what poor Lady Mondegreen did for someone to slay her."[1] "Sadly," the article concludes, "*mondegreen* does not appear in the Bank of English – yet."

In substance, the general argument of this paper is that syllable structure may help listeners avoid mondegreening every sentence they hear. More specifically, we argue that syllable onsets are generally highly salient in the signal, and serve as potential alignment points for the lexical search process. The Syllable Onset Segmentation Hypothesis (SOSH) bears some close similarity to the Metrical Segmentation Strategy (MSS) proposed for English [2] and contrasts with alternative proposals [3] which assume that the signal is recoded and categorized prelexically in terms of syllable-sized units mediating lexical access. We assume that the lexical mapping process is based on smaller-size units (phonemes or features), but that syllable structure determines privileged alignment points.

We present an overview of three sets of studies bearing on explicit syllabification, on the influence of syllable/word misalignments on online measures of lexical access, and on resyllabification at word boundaries.

2. EXPLICIT SYLLABIFICATION

One data source about segmentation comes from explicit syllabification tasks. In a study investigating adults' syllabification [4], American subjects were required to reverse the syllables in bisyllabic words with single intervocalic consonants. The listeners' performance was not consistent, the intervocalic consonant being sometimes placed in the first syllable (*melon* reversed as *lon-me*), and sometimes in the second (*melon* > *on-mel*). In addition, there were also many *ambisyllabic* responses where the consonant was placed in both syllables (*lon-mel*). The more sonorous the consonant, the greater the probability that it was placed in the first syllable rather than in the second syllable, and that it would elicit ambisyllabic responses.

Some years ago, we launched a similar investigation of syllabification in French [5]. Our first study aimed at assessing the syllabification of singleton intervocalic consonants, and followed closely Treiman & Danis. Much to our surprise, we observed that as the American participants, French listeners produced about 15% of ambisyllabic responses. We reasoned that such unexpected responses could occur if decisions about syllable onsets (henceforth marked by an opening square bracket, "[") and syllable offsets ("]") involve distinct processes. Follow-up experiments used slightly different tasks, in which different groups of participants had to repeat either the first or the second part of the same stimulus words. As shown in Figure 1, the vast majority of second-part responses included the intervocalic consonant (e.g. *ballon* > *lon*, i.e.[CV), as predicted by all phonological analyses of French. But first-part responses split nearly evenly between CV and CVC (*ballon* > *ba* or *bal*). Moreover, the tendency to produce closed (CVC) first-part responses was more manifest for more sonorous consonants, and also influenced by orthographic gemination (e.g. *ballon* vs *palais*).

In another study [6], we examined the influence of orthography more directly by comparing syllabification preferences in five-year-old nonreaders and ten-year-old literate children, using a task, design and materials similar to the previous adult experiment. In the readers group, CVC first-part responses were more frequent for orthographically geminated words than for words with a single orthographic consonant. No such spelling effect emerged for prereaders. Otherwise, in both age groups, the pattern of responses closely replicated the adult data regarding the dissociation between first and second-part responses, and the specific effect of sonority on first-part responses. Thus, it does not seem that this dissociation is due to the influence of literacy acquisition.

We also examined syllabification by American English speakers in Brussels. The experiment and materials were designed exactly as in the French study described above, but we also contrasted first and second syllable-stressed words. For second-syllable stressed stimuli, the pattern closely resembles the French data: more than 90% of CV[CV responses for the onset of the

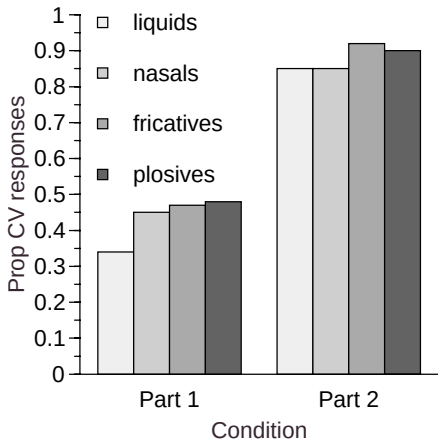

Figure 1. Proportion of canonical syllabification responses (CV]CV for Part 1 and CV[CV for Part 2).

second part, and a lot of variability for the offset of the first part. For the first-syllable stress words however, a different pattern was observed, suggesting a CVC][V syllabification: First-syllable responses nearly always included the intervocalic consonant and about 30% of second-syllable responses for words with a single orthographic consonant did not include the consonant.

Overall, these findings indicate that French listeners are not consistent in their segmentation of syllables, specially for syllable offsets. Interestingly, by separating those responses that required the determination of the syllable onsets from those requiring decisions for syllable offsets, we observed a clear dissociation. The former were more consistent than the latter; second syllable responses nearly always began with the consonant, whereas responses for the first syllable varied more, often including the intervocalic consonant.

3. ON WORD-SYLLABLE MISALIGNMENT

We take the syllabification data to indicate that syllable onsets constitute reliable segmentation points in the signal, and we hypothesize that syllable onsets are used as privileged alignment points for lexical search in continuous speech recognition. In further studies, we have tried to address this issue by investigating the processing of words embedded in multisyllabic carriers. We examined whether a misalignment between a syllable onset and the target word onset delays the recognition of the word.

3.1. Word Spotting

Some empirical support for SOSH comes from a word spotting experiment [7] in which participants made speeded manual responses when they detected monosyllabic words embedded at the beginning or end of bisyllabic nonwords (e.g., *lac* in *lactuf* or *zunlac*). All target words began with a liquid and ended with an occlusive. Syllable structure was manipulated by varying the consonant immediately preceding or following the target word. Thus, our manipulation relied on generally accepted phonotactic principles for French consonant clusters, which state that obstruent-liquid clusters are tautosyllabic (as in *la.cluf* or *zu.glac*), whereas obstruent-

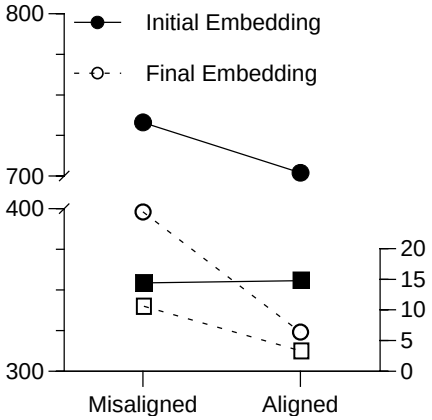

Figure 2. Word spotting RTs (circles) and Errors (squares).

obstruent, liquid-liquid and nasal-liquid clusters are separated (as in *lac.tuf* or *zun.lac*). All consonant clusters used were attested in French polysyllabic words. Two groups were tested, one with the onset alignment and the other with the offset alignment manipulation.

According to SOSH, the lexical cost due to word/syllable misalignment should be greater for final embedding, which corresponds to onset misalignment, than for initial embedding. As shown in Figure 2, this prediction was confirmed both by RT and error data. Onset misalignment (*zu[glac* vs. *zun[lac*) induced significant RT (74 msec) and error (7.3%) effects, whereas offset misalignment (*la]cluf* vs *lac]tuf*) revealed only a small and non-significant RT effect.

3.2 Crossmodal Repetition Priming

Another study conducted by Gregory Leclercq in Brussels used cross-modal repetition priming. Participants had to perform a lexical decision task on visually presented monosyllabic words, preceded by a related or unrelated bisyllabic auditory nonword. Visual target items were presented for 50 msec, and synchronized with the offsets of the auditory stimulus.

In the related pairs, the auditory prime word was embedded in a bisyllabic nonword carrier. We contrasted four related conditions, in a design similar to the previous word-spotting experiment. Prime words were embedded either at the beginning or at the end of the carriers, and syllable structure was manipulated by varying the nature of the consonant immediately preceding or following the target word. Thus *nam.robe*, *ja.vrobe*, *rob.jaf* and *ro.blane* were the four possible related primes for the visual target ROBE. Different groups of participants were tested with initial and final prime embedding positions.

Based on SOSH, we predicted that listeners should show greater sensitivity to onset (final embedding, *nam.robe* vs *ja.vrobe*) than to offset misalignment. A first analysis of the results did not seem to confirm our expectations. As shown in Figure 3, when the data were collapsed for both blocks together, no hint of the predicted interaction was observed. In fact, a significant relatedness by alignment interaction was found, but it did not vary as a function of position. However a different

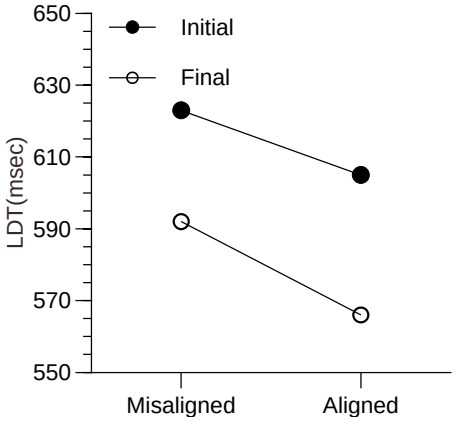

Figure 3. Mean Lexical Decision Times in the Related Conditions, as function of the position of the embedded prime word in the carrier, and of syllable structure.

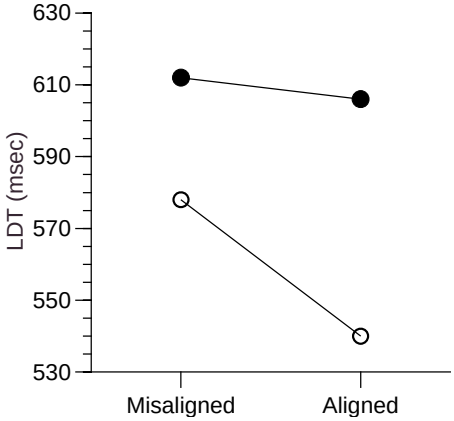

Figure 4. Mean Lexical Decision Times for first target presentation (Block 1) in the Related Conditions.

picture emerged when we looked at the data for Block 1 only. As shown in Figure 4, results for Block 1 conform exactly to our predictions: A significant alignment effect was only observed for the final embedding condition. Although these data are clearly preliminary, and further testing is needed to confirm the outcome, they provide at least encouraging support for our hypothesis.

4. ON RESYLLABIFICATION

Obviously the interest of SOSH for continuous word recognition depends on the likelihood that word and syllable onsets coincide. Paradoxically, given the processing costs observed in the previous experiments, SOSH would be a hindrance rather than an aid if misalignments are the rule in continuous speech.

One reason such misalignment might be frequent is resyllabification. Since syllables pertain to the domain of suprasegmental structure, it is generally admitted that syllable reorganisations occur at word junctures (as illustrated by the "laid him on" / "Lady Mon" anecdote). Accordingly, some models of speech production [8] propose that phonetic encoding involves syllabic articulatory plans computed or accessed after word boundaries have been erased and resyllabification has occurred.

It thus seemed critical to assess whether such phonetic resyllabification does indeed occur. To that end, we constructed pairs of two-word phrases that had two syllables in common but differed by the position of the word boundary (e.g. *tant#rou* in *une tante roublarde*, *tan#trou* in *des temps troublants*, see [9]). Across various experiments, the consonant cluster at the word juncture was either obstruent+liquid (OBLI) or /s/+obstruent (SOB). These materials were used to investigate several related issues: (1) whether the lexical intent of speakers determines systematic phonetic variations; (2) whether listeners are sensitive to such variation in a syllabification task; and (3) whether it affects lexical processing.

In the production study, naive French speakers produced pairs of such two-word phrases, either in isolation, or in sentence context. Durational analyses showed systematic lengthening of the pre-boundary vowel and of the liquid consonant in CVC#CV sequences. In contrast, no reliable effects were obtained for SOB clusters.

For the perceptual studies, the shared bisyllabic sequences (e.g. *tantrou*) were extracted from the OBLI noun phrases. In one study, participants were asked to repeat either the first or second syllable of the sequences under time pressure, and we examined the pattern of syllabification. As shown in Figure 5, the proportion of canonical syllabification responses (CV)[CCV) is much higher for the second syllable repetition condition than for the first syllable repetition condition, thus replicating the findings described above. In addition, and more importantly, syllabification responses indeed reflect the manipulation of word boundary position.

In another study [9], the same stimuli were used in a word spotting task. Participants detected CVC words either at the beginning (*tante* in *tantrou*) or at the end (e.g. *roche* in *icroche*) of nonsense bisyllables. Whereas no effect was observed for SOB clusters, significant phonetic misalignment effects occurred for OBLI clusters: Listeners were about 100 msec faster to detect *roche* in *i#croche* than in *ic#roche*; and they were 140 msec faster to detect *tante* in *tant#rou* than in *tan#trou*. The finding of an offset misalignment effect in the latter case appears to run counter the prediction of SOSH. However, other data [10] show that this effect is caused by competition related to the lexical status of the second syllable.

In sum, these results run counter the view that in production syllable structure is superimposed on the phonological string independently of word juncture information. At least for OBLI clusters, we found that systematic phonetic variations occur at word boundaries and that listeners are indeed sensitive to them. The absence of any word boundary cues in SOB clusters is intriguing. However, it fits well with other syllabification and production data [11, 12] suggesting that /s/ has a special status as regards syllabification or can be considered extra-syllabic in such sequences.

In the course of our examination of resyllabification, we have started examining in detail the phonetics of word boundary phonemes. One might wonder whether the phonetic cues that we have identified subserve syllabification, or constitute independent cues to lexical segmentation. One radical interpretation of our results is that resyllabification does not occur: e.g. *tan#trou* is realised *tan.trou* and *tant#rou* is realised *tant.rou*:

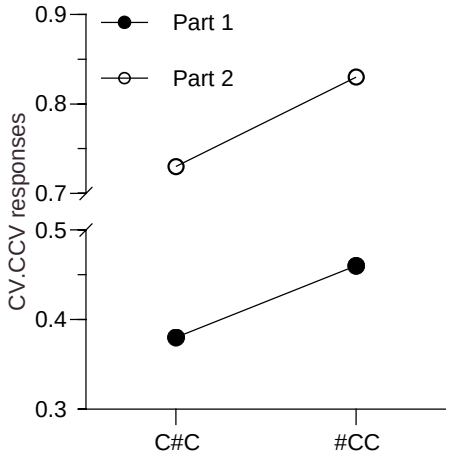

Figure 5. The effect of word boundaries on syllabification.

perceptually, phonetic cues inform syllabification, and syllabification informs lexical segmentation. In this perspective, the absence of phonetic variation in SOB clusters simply means that the two potential syllabification of SOB clusters are not phonetically marked, e.g. *ra.stu* does not differ phonetically from *ras.tu*. Of course, it remains to be determined whether phonetic cues distinguish *Lady Mon* from *laid him on.*

5. CONCLUSIONS

SOSH is based on two ideas: onsets, relative to offsets constitute more reliable segmentation points; syllable boundaries tend to coincide with word boundaries. Regarding the first claim, there is much phonological evidence that indicates that word/syllable initial consonant are more salient, more stable both in terms of their phonetic characteristics but also in terms of language change. With respect to the second claim, the correspondence of word and syllable onsets has *prima facie* value. The notion of syllable has a long history in phonology and phonetics. It is one of those ubiquitous notions that everyone endorses but no one can define. Nevertheless, some of the major theories define syllable boundaries on phonotactic grounds, that is, from attested word onsets and offsets. While the phonotactic perspective may not capture all the phonological facts, it definitely suggests that overall, word and syllable boundaries should correspond, except for special cases such as liaison, or enchaînement. The frequency and phonetic characteristics of such phenomena require further examination

Perhaps the most interesting implication of SOSH concerns the issue of language specificity. Indeed the similarity between our results and other findings in Dutch and English, and the close relation between SOSH and MSS suggests that the syllable's role might be more similar across languages than has generally been admitted. The universal status of SOSH is thus open to investigation. Another issue is how syllable structure combines with other bottom-up and top-down segmentation cues. We have started to examine how SOSH and competition interact [10]. It remains for further research to determine the respective roles of syllable structure, rhythmic and prosodic cues, and word knowledge in lexical segmentation.

While further supporting evidence would certainly help the cause of SOSH, we would argue that it constitutes a more plausible and more viable hypothesis than the prelexical syllabic classification hypothesis, and that it opens interesting avenues for further research.

6. ACKNOWLEDGEMENTS

This research is supported by grants from the Communauté française de Belgique (A.R.C. 96/01-203) and from the Swiss FNRS (Project 1113-049698.96). We thank Cécile Fougeron for her comments and many useful discussions. Nicolas Dumay is research assistant at the Belgian FNRS. Correspondence to Alain Content, LAPSE—ULB CP191, Avenue F. D. Roosevelt, 50, B–1050 Bruxelles. Email: Alain.Content@ulb.ac.be.

7. REFERENCES

[1] Wordwatch (2000). Firstable and other mondegreens, http://titania.cobuild.collins.co.uk/cgi-bin/wwatchlook.

[2] Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology : Human Perception and Performance*, **14**, 113-121.

[3] Segui, J., Dupoux, E., & Mehler, J. (1990). The role of the syllable in speech segmentation, phoneme identification and lexical access. In G. Altmann (Ed.), *Cognitive Models of Speech Processing*, MIT Press, 263-280.

[4] Treiman, R., & Danis, C. (1988). Syllabification of intervocalic consonants. *Journal of Memory and Language*, **27**, 87-104.

[5] Content, A., Kearns, R., & Frauenfelder, U. (1999). Boundaries versus onsets in syllabic segmentation. *Ms submitted for publication*.

[6] Content, A., Frauenfelder, U., & Dumay, N. (1999). Segmentation syllabique chez l'enfant, *Journées d'Etudes Linguistiques - La syllabe.* Université de Nantes, 69-74.

[7] Dumay, N., Banel, M. H., Frauenfelder, U., & Content, A. (1998). Le rôle de la syllabe: segmentation lexicale ou classification? *Actes des XXIIèmes journées d'étude sur la parole*, 33-36.

[8] Levelt, W., & Wheeldon, L. (1994) Do speakers have access to a mental syllabary? *Cognition*, **50**, 239-269.

[9] Dumay, N., Frauenfelder, U., & Content, A. (1999). Acoustic-phonetic cues to word boundary location: evidence from word spotting, *Proceedings of ICPhS*, San Francisco, CA, 281-284.

[10] Dumay, N., Frauenfelder, U., & Content, A. (2000). Acoustic-phonetic cues and lexical competition in segmentation of continuous speech. *Proceedings of SWAP*, Nijmegen:MPI for Psycholinguistics.

[11] Rialland, A. (1994). The phonology and phonetics of extrasyllabicity in French. In P. A. Keating (Ed.), *Phonological structure and phonetic form: Papers in Laboratory Phonology* (Vol. 3), 136-159.

[12] Treiman, R., Gross, J., & Cwikiel-Glavin, A. (1992). The syllabification of /s/ clusters in English. *Journal of Phonetics*, **20**, 383-402.