IF SYLLABLES WERE CLASSIFICATION UNITS IN SPEECH PERCEPTION, AUDITORY PRIMING WOULD SHOW IT

Nicolas Dumay [1], Alain Content [2] and Monique Radeau [2,3]

[1] University of Bristol, United Kingdom; [2] Université libre de Bruxelles, Belgium; [3] National Fund for Scientific Research, Belgium

n.dumay@bristol.ac.uk; acontent@ulb.ac.be; moradeau@ulb.ac.be

ABSTRACT

Two auditory priming experiments tested whether the final overlap effect relies on syllabic representations. Amount of shared phonetic information and syllabic correspondence between prime and target nonwords were varied orthogonally. In the related conditions, CV.CCVC primes and targets shared the last syllable (e.g. *pinclude-viclude*) or the last syllable minus one phoneme (e.g. *pinclude-viflude*); conversely, CVC.CVC primes and targets shared the last syllable (e.g. *goltibe-purtibe*) or the last syllable plus one phoneme (e.g. *goltibe-pultibe*). Both experiments required to repeat back the targets, with Experiment 2 including foils. The facilitation induced by related primes increased with the number of shared phonemes, and was by and large independent of syllabic correspondence. There is thus little evidence for pre-lexical syllabic classification.

Keywords: speech perception; auditory priming; syllable; shadowing

1. INTRODUCTION

Since the study by Mehler and colleagues [1] showing that French listeners are faster at detecting /ba/ in *ba.lance* than in *bal.con*, and vice versa for /bal/, the idea that syllables are classification units in speech perception, at least in the Romance languages, has been quite popular. According to this hypothesis, speech would be categorized in syllable chunks, which would then be used to both access the lexicon and derive the identity of the phonetic segments [2]. Despite the fact that the interpretation of the so-called "syllabic effect" has been repeatedly called into question (e.g. [3,4]), the heightened sensitivity of Romance language native speakers to the syllabification of speech has been established using other paradigms, including attentional allocation, phonological migrations, word spotting and crossmodal priming (see [5] for an overview). For instance, in Italian, the latter paradigm showed that [si.l], fragment compatible with *si.lenzio*, evokes the related word *rumore*, whereas the incompatible fragment [sil.] does not, and vice versa for [sol] and a word like *sol.dato* [6]. However, if these results are compatible with the syllabic classification hypothesis, they do not constitute unambiguous evidence in its favor, either because the effects are more likely to tap into segmentation (rather than classification) mechanisms, or because alternative classification units, for instance, positional allophones, could also explain the data.

A phenomenon which appears perfectly suited to test directly the idea of syllabic classification is phonological priming, whereby the processing of a stimulus (the target) is affected by the prior presentation of another stimulus (the prime), if they contain similar segments. Presumably, this effect comes from the necessity for the listener to evoke twice the same representation within a short period of time. Studies on this phenomenon show a neat dissociation depending on the location of the common segments. In tasks such as target repetition and lexical decision, initial overlap (e.g. *cube-cure*) tends to produce inhibition [7,8], reflecting lexical competition. In contrast, final overlap (e.g. *ramp-lamp*) systematically induces facilitation [9-11]. Beyond its transient nature, properties of the final overlap facilitation suggest that it taps into processing levels independent of the lexicon, most likely prelexical (see [9]). Indeed, unlike initial overlap inhibition, it requires that both prime and target be auditory; it does not depend on the lexical status of either the prime or the target, nor on the lexical status of the shared portion; and it is insensitive to the prime-target relative frequency.

Given the likely prelexical nature of the initial overlap priming effect, the goal of this research was to examine whether it relies on syllabic representations. Two experiments orthogonally varied the amount of shared phonetic information

and the syllabic correspondence between prime and target nonwords. The shared sequence was always a nonword. In the related conditions, CV.CCVC primes and targets shared the last syllable or the last syllable minus one phoneme; conversely, CVC.CVC primes and targets shared the last syllable or the last syllable plus one phoneme. Both experiments used target repetition, with Experiment 2 including foils to discourage participants from completing targets by the ending segments of the primes. To determine whether strategies (expected to induce a cost on unrelated targets) were involved, each experiment was compared to a baseline session including only unrelated trials. If the final overlap effect is due to the activation of syllabic representations, no effect should be obtained in the infrasyllabic condition, and the syllabic and suprasyllabic conditions should produce equal facilitation. In contrast, if the effect simply depends on the amount of shared phonetic information, it should increase with the number of shared phonemes, independently of syllabic structure.

2. EXPERIMENT 1

2.1. Methods

2.1.1. Participants

Fourty-eight native French speakers with no auditory or language impairment were tested.

2.1.2. Materials, design and procedure

The key-materials were 120 quadruplets composed of three primes and one target. All stimuli were CVCCVC phonotactically legal nonwords in French, such that their CCVC and CVC final portions never corresponded to a word. Nonwords in half the quadruplets had an OBstruent-LIquid (OBLI) consonant cluster, so that the syllable boundary was located just before the cluster (CV.CCVC). In the other half, a LIquid-OBstruent (LIOB) cluster was used, so that the syllable boundary was located between the consonants (CVC.CVC). Within each quadruplet, two primes rhymed with the target (e.g. *pinclude*), by sharing with it either the four or the three final phonemes (e.g. *viclude* vs. *viflude*), whereas the third (e.g. *vifroge*), phonologically unrelated to the target, served as a control. Hence, the four-phoneme overlap corresponded to the final syllable in CV.CCVC items but was larger than the final syllable in CVC.CVC ones. Conversely, the three-phoneme overlap was one-phoneme smaller than the final syllable in CV.CCVC items but corresponded to the final syllable in CVC.CVC ones. Fillers were 200 pairs of phonologically unrelated CVCCVC items. To make sure that participants would engage in lexical processing, 160 had a word as target. All filler primes were nonwords. Within each pair, prime and target had the same syllabication, half the word and nonword target pairs comprising .CC items, while the other pairs comprised C.C items.

All stimuli were produced in a sound-proof booth by a male native speaker of French and digitized at 44.1 kHz/16-bit using SoundDesigner. CV.CCVC stimuli were shorter than CVC.CVC ones, but for both syllabic structures primes were closely matched in duration.

For the primed version, prime-target pairs were distributed over three lists of 320 trials, of which 25% were related pairs. Targets appeared once in each list, but were each time preceded by a different prime to counterbalance overlap across lists. Targets positions were identical across the lists, which were split into five even blocks to track performance on the unrelated trials. Trial order was such that no more than two related trials occurred in direct succession. A fourth list comprising only unrelated trials was created for the baseline version.

Participants were required to repeat the second item of each spoken pair presented via headphones. Targets began 20 ms after prime offset. Response onsets were detected via a headset microphone interfaced with the computer. Onset detection and accuracy were checked online. Twelve participants were assigned to each list. Before the test, they practiced the task on 40 unrelated trials. Block order was rotated across participants to avoid item order effect.

2.2. Results and discussion

Latency analyses were based on correct responses RTs, excluding those (0.8%) shorter than 150 ms or longer than 1500 ms, or misaligned with response onset. Three targets repeated incorrectly more than 1/3 of the time were also discarded from all analyses. As attested by the effect of overlap ($F1(2,70) = 33.74$, $p < .001$; $F2(2,230) = 28.10$, $p < .001$), and further planned comparisons, latencies were faster overall in the four-phoneme (648) than in the three-phoneme condition (657; $ps \leq .053$), and faster in the latter than in the

control condition (684; *ps* < .001; see Table 1). The crucial syllabic structure x overlap interaction was significant only by participants (*F1*(2,70) = 3.31, *p* < .05; *F2* < 1). It resulted from the infrasyllabic condition inducing a significant but smaller facilitation (*ps* < .05) than the syllabic and suprasyllabic conditions, no different from each other (*ps* > .13). However, contrary to the syllabic hypothesis, the difference in facilitation between the three-phoneme and four-phoneme overlap conditions in .CC items was no larger than its C.C counterpart (*p* > .3).

		Overlap			Ctrl vs.	
	σ	Ctrl	Three	Four	Three	Four
RTs(ms)	.CC	669	650	638	19	31
	C.C	699	664	658	34	41
Ers(%)	.CC	8.8	3.8	2.9	5.0	5.8
	C.C	6.4	2.5	3.2	3.9	3.2

Table 1: Mean reaction times (RTs), error rates (Ers) and priming effects (relative to Ctrl) as a function of syllabic structure and overlap for Experiment 1.

Error rates were also sensitive to overlap (*F1*(2,70) = 22.02, *p* < .001; *F2*(2,230) = 19.9, *p* < .001), and this independently of syllabic structure (*F1*(2,70) = 1.96, *p* > .1; *F2*(2,230) = 1.28; *p* > .25). The three- and the four-phoneme condition did not differ from each other (*Fs* < 1), but were (4.5%) less error-prone than the control condition (*ps* < .001). As shown by the time-window analysis comparing the primed and the baseline version (cf. Figure 1), the profile of RTs to controls did not show any processing cost in the presence of related trials (*F* < 1), which argues against the idea that the above priming effects have a strategic component.

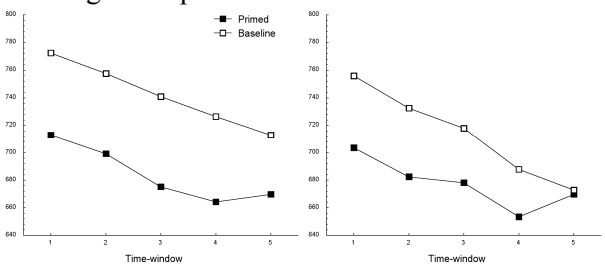

Figure 1: Evolution of mean RT to unrelated trials over five successive time-windows, for the primed and baseline versions of Experiments 1 (left) and 2 (right).

3. EXPERIMENT 2

Experiment 2 included foils meant to discourage participants from anticipating the final segments of the targets on the basis of those of the primes. It also allowed us to assess the viability of the syllabic structure by overlap interaction trend found in Experiment 1.

3.1. Methods

3.1.1. Participants

Another 48 native French speakers were tested.

3.1.2. Materials, design and procedure

Materials were identical to Experiment 1, except that 24 filler pairs (12 per syllabic structure) were replaced by foils, introduced at regular intervals. Foils were such that primes and targets (seemingly) rhymed as of the first or second medial consonant, but then diverged upon one of the next segments (e.g. *vuglète-bouglèfe*).

3.2. Results and discussion

		Overlap			Ctrl vs.	
	σ	Ctrl	Three	Four	Three	Four
RTs(ms)	.CC	669	652	641	16	28
	C.C	685	669	660	16	24
Ers(%)	.CC	5.6	4.2	2.8	1.4	2.8
	C.C	4.8	2.3	2.5	2.5	2.3

Table 2: Mean RTs, error rates and priming effects as a function of syllabic structure and overlap for Experiment 2.

Using the rejection criteria of Experiment 1, 0.7% of the RTs were discarded. As shown by the effect of overlap (*F1*(2,70) = 23.44, *p* < .001; *F2*(2,230) = 12.46, *p* < .001), and planned comparisons, latencies were 10 ms faster in the four-phoneme than in the three-phoneme condition (*ps* < .05), and also 16 ms faster in the three-phoneme than in the control condition (*ps* < .004). However, in contrast to Experiment 1, there was no hint of an interaction between syllabic structure and overlap, i.e., the strength of the facilitation was exclusively determined by the number of shared phonemes (see Table 2). As in Experiment 1, the related conditions were less error-prone than the control condition (here, by 2.3%; *ps* < .05) and did not differ from each other. Again, there was a significant effect of overlap (*F1*(2,70) = 6.90, *p*< .002; *F2*(2,230) = 7.17, *p* < .001), but clearly no interaction between overlap and syllabic structure.

Concerning the evolution of the RTs to control targets (see Figure 1), there was an interaction between time-window and list version (*F*(4,112) = 3.35, *p* < .05). However, it was due to a change in

performance only at block 5, which on its own, and given the slightly smaller overlap effects in the second than in the first experiment, is unlikely to explain the observed priming.

As shown by combined analyses of the two primed versions, the presence of foils had no reliable influence on the pattern of RTs, with no experiment by overlap interaction ($F1(2,140) = 1.61$, $p > .2$; $F2(2,230) = 2.68$, $p = .072$) or three-way interaction with syllabic structure ($F1(2,140) = 2.14$, $p > .12$; $F2 < 1$), but it induced a significant 2.4% reduction on the errors to control targets ($ps < .05$). Note that accuracy in the related conditions was virtually at ceiling already without foils. Crucially, with both experiments taken together, there was no hint of an interaction between syllabic structure and overlap either on RTs or on errors (all $Fs \leq 1.68$, $ps > .15$).

The use of foils was motivated by the fact that participants could have engaged (voluntarily or not) in completing targets by the ending segments of the primes. This predicts that errors with contamination by the second syllable of the prime (e.g. *gulṭcaise* for *gulcaise* when preceded by *pourṭode*) should be proportionally larger in presence than in absence of related trials. Against the above interpretation, but in line with both the absence of an effect of foils and the time-window analyses, such errors were (statistically) no more frequent in the primed than in the baseline versions (Exp 1: 20.6% vs. 15.6%; Exp 2: 13.7% vs. 9.5%).

4. GENERAL DISCUSSION

The syllabic classification hypothesis refers to the idea that listeners would recode speech into structured representations corresponding to the syllables present in the signal in order to access lexical memory [2]. If this hypothesis is correct, changes in the way listeners process speech should be observed whenever two syllables temporally close to one another require computation or activation of the very same mental representation. Using the (facilitation) effect of final overlap, which is very likely to tap into prelexical levels of processing, we found no reliable evidence of syllabic priming in a repetition task. In both experiments, significant facilitation was observed for the crucial infrasyllabic overlap (.CCVC). This, in itself, allows to reject the strongest "stage approach" of the syllabic hypothesis, in which no other units than syllables would be used in decoding speech.

However, it could still be the case that syllables are only one level of representation among several, in a cascaded model. Against this view, which could account for the infrasyllabic facilitation observed, we found no reliable modulation of the facilitation effect by the syllabic nature of the overlap: simply, the more phonemes primes and targets had in common, the larger the facilitation. It remains to be determined whether these effects are phonological in nature or on the contrary rely on some sort of acoustical trace.

In view of these data, it seems to us that the alternative interpretations available to account for syllabic influences, in particular in crossmodal semantic priming and phonological migrations should be favored. The syllable may not be a unit of speech perception; however, as a prosodic unit, its impact on the realization of segments is such that the syllabic structure has yet an indirect influence on the lexical mapping and alignment processes [3-6].

5. REFERENCES

[1] Mehler, J., Dommergues, J.Y., Frauenfelder, U.H., Seguí, J. 1981. The syllable's role in speech segmentation. *J. Verb. Learn. Verb. Behav.* 20, 298-305.

[2] Mehler, J., Dupoux, E., Seguí, J. 1990. Constraining models of lexical access: The onset of word recognition. In: Altmann, G.T.M. (ed.), *Cognitive models of speech processing*. Cambridge, MA: MIT Press, pp. 236-262.

[3] Kolinsky, R. 1998. Spoken word recognition: A stage-processing approach to language differences. *Eur. J. Cogn. Psychol.* 10, 1-40.

[4] Content, A., Meunier, C., Kearns, R.K., Frauenfelder, U.H. 2001. Sequence detection in pseudowords in French: Where is the syllable effect? *Lang. Cogn. Proc.* 16, 609-636.

[5] Dumay, N., Frauenfelder, U.H., Content, A. 2002. The role of the syllable in lexical segmentation in French: Word-spotting data. *Brain Lang.* 81, 144-161.

[6] Tabossi, P., Collina, S., Mazetti, M., Zoppello, M. 2000. Syllables in the processing of spoken Italian. *J. Exp. Psychol. Hum. Percept. Perf.* 26, 758-775.

[7] Radeau, M., Morais, J., Dewier, A. 1989. Phonological priming in spoken word recognition: Task effects. *Mem. and Cogn.* 17, 525-535.

[8] Dufour, S., Peereman, R. 2003. Inhibitory priming effects in auditory word recognition: When the target's competitors conflict with the prime word. *Cognition* 88, B33-B44.

[9] Dumay, N., Benraïss, A., Barriol, B., Colin, C., Radeau, M., Besson, M. 2001. Behavioral and electrophysiological study of phonological priming between bisyllabic spoken words. *J. of Cogn. Neurosc.* 13, 121-143.

[10] Norris, D., McQueen, J.M., Cutler, A. 2002. Bias effects in facilitatory phonological priming. *Mem. and Cogn.* 30, 399-411.

[11] Chéreau, C., Gaskell, M.G., Dumay, N. 2007. Reading spoken words: Orthographic effects in auditory priming. *Cognition* 102, 341-360.