Proceedings of the Workshop on Spoken Word Access Processes, 29-31st May 2000, Max-Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

ACOUSTIC-PHONETIC CUES AND LEXICAL COMPETITION IN SEGMENTATION OF CONTINUOUS SPEECH

Nicolas Dumay,
Laboratoire de Psychologie Expérimentale, Université libre de Bruxelles, and F.N.R.S., Belgium
Uli H. Frauenfelder
Laboratoire de Psycholinguistique Expérimentale, Université de Genève, Switzerland
and **Alain Content**
Laboratoire de Psychologie Expérimentale, Université libre de Bruxelles, Belgium

ABSTRACT

The present research examined the interplay between lexical competition and acoustic-phonetic cues in word segmentation and recognition. Lexically ambiguous bisyllabic carriers were used in word-spotting experiments that required the participants to detect CVC or CV initially embedded words. The syllabification of the medial cluster (C.C vs .CC) and the lexical status of the post-boundary final chunk were manipulated. The word-spotting responses to CVC words were clearly inhibited by the overlapping word, leading to a target-offset misalignment effect. The CV word-spotting latencies also showed a misalignment effect, that tended to be reduced when the target was followed by a word. These results are interpreted in terms of a framework which combines the PWC and a prelexical segmentation heuristic based upon the onsets of syllables.

1. INTRODUCTION

Speakers do not provide listeners with systematic word-boundary markers in the signal. Nonetheless, listeners generally manage to segment the message into its meaningful constituent parts by exploiting various phonological and prosodic regularities of the spoken form. According to one useful heuristic, listeners assume that certain prosodic boundaries correspond to word boundaries. Indeed, for stress-timed languages like English and Dutch, it has been proposed that listeners apply the Metrical Segmentation Strategy (MSS) and rely on foot boundaries or more precisely, on the strong syllables to initiate lexical search [1].

The objective of this paper is to provide support for a comparable heuristic that we believe to be used by native speakers of French, which has different prosodic properties. French is generally described as being syllable-timed. Stress is not lexically contrastive, primary stress falls on the final syllable of accentual groups, and there is no vocalic reduction. We have proposed the Syllable Onset Segmentation Heuristic (SOSH) for French ([2],[3]) that is quite similar to the MSS described above. According to SOSH, unlike for English and Dutch, every French syllable onset constitutes an alignment point for lexical search.

Evidence for such a Syllable Onset Segmentation Heuristic comes from word-spotting experiments [4], in which the alignment between word and syllable boundaries was manipulated. Consistent with the SOSH hypothesis, recognition of the embedded word (e.g. lak) showed a greater processing cost for onset misalignment (zyn.lak vs zy.glak) than for offset misalignment (lak.tyf vs la.klyf). Interestingly, analogous word-spotting studies in Dutch [5] and English [6] that manipulated either syllable alignment or positional allophones also reported larger effects at word onset than at word offset.

However, the difference in perceptual cost between onset and offset alignment was not replicated in another set of experiments [7]. These experiments aimed at assessing the role of acoustic-phonetic syllabification cues on word-spotting performance, in obstruent-liquid clusters. Participants spotted CVC words embedded initially or finally, in <u>CVC</u>CV (e.g. tɑ̃t in tɑ̃tʀu) or VC<u>CVC</u> (ʀɔʃ in ikʀɔʃ) nonce strings. The carrier strings were extracted from paired two-word phrases that differed in the word-boundary location, either within the cluster (i.e. tɑ̃t#ʀu from *tante roublarde* and ik#ʀɔʃ from *magique roche*) or before the cluster (i.e. tɑ̃#tʀu from *temps troublant* and i#kʀɔʃ from *demi-croche*). The results showed that offset and onset misalignments between targets and word/syllable boundaries produced processing costs of the same magnitude.

These results appear to be inconsistent with the findings reported above and suggest that contrary to the SOSH hypothesis, offset alignment cues are as important as onset alignment cues. However, another possible explanation for these effects is in terms of lexical competition. Indeed, several studies ([8],[9]) demonstrated that the activation of an initially embedded word was negatively affected by the number of candidates starting with the final consonant of the target (cf. [8] *melk* in *melkaam* vs *melkeum*), whereas no such effect was observed when the target and candidate set did not overlap (*bel* in *belkaam* vs *belkeum*).

In our experiment [7], the carriers were devised so that the post-boundary portions (i.e. tʀu and ʀu, in tɑ̃tʀu) were always words in order to prevent lexical biases from favoring one particular segmentation. Thus, there might be competition between the target *tɑ̃t* and the CCV word *tʀu* since they share the pivotal consonant. In the present study, we aimed, through the word-spotting task, to determine whether the emergence of offset alignment effects actually depends upon the lexical activation of following portions of input. According to

SOSH, the overlapping final word *tʀu* should be more activated in *tã#tʀu*, in which it is aligned with a syllable onset, than in *tãt#ʀu*. Hence, its inhibitory influence on the initially embedded target *tãt* should be enhanced. This inhibitory effect should disappear if the final portion of carriers is neither a word nor the beginning of longer candidates. Experiment I determined the role of subsequent overlapping competitors, using CVCs as targets.

2. EXPERIMENT I

This experiment assessed how the effect of offset misalignment between words and the acoustic-phonetic structure is dependent upon or enhanced by the activation of a target-overlapping competitor. CVC initially embedded words had to be spotted in CVCCV(C) lexically ambiguous strings, in which both the acoustic-phonetic syllabification cues, and the lexical status of the CCV(C) final portion were orthogonally manipulated.

2.1. Method

2.1.1. Participants

Forty-eight students participated in the experiment. All were native speakers of French, and reported no hearing or speech disorders.

2.1.2. Materials

Seventy-two CVC (or CCVC[1]) French words (e.g. *tante* [tãt]) were selected as targets. The final consonant of these words was either a plosive or a fricative. In all targets, the initial CV portion was also a word (*temps* [tã]). Each target appeared at the beginning of two types of nonce bisyllabic strings with an obstruent-liquid medial consonant cluster (pʀ, tʀ, fʀ, bʀ, dʀ, vʀ, kl, gl).

To examine the combined effects of offset misalignment and lexical competition, targets were embedded in **<u>CVC</u>**CV carriers (e.g. tãtʀu), such that their CCV final portion was a word, and thus, a target-overlapping competitor (i.e. tʀu). In these carriers, the CV final portion was a word too. Two acoustic-phonetic versions of each carrier were obtained. These were extracted from the beginning of paired longer trisyllabic two-word phrases that differed in the word-boundary location (i.e. C#C... as from *tante roublarde*, or #CC... as from *temps troublant*). Thus, in the <u>aligned</u> condition, the offset of the target word matched the intended word boundary, while the onset of the CCV target competitor did not (tãt#ʀu). In contrast, in the <u>misaligned</u> condition, the onset of the competitor matched the intended word boundary, while the offset of the target word did not (tã#tʀu).

To examine the effect of misalignment without lexical competition, targets were embedded in carriers (e.g. tãtʀøm), devised with the constraints that (1) the target-following liquid was the same as in the competition situation, (2) their CCVC final portion was neither a word nor the beginning of a longer one, and (3) the CVC final portion was neither a word nor the beginning of a longer one. Again, two versions of each carrier were obtained, extracted from paired trisyllabic two-component phrases made-up of a monosyllabic word and a bisyllabic pseudoword (i.e. C#C... as from *tante reumlite*, or #CC... as from *temps treumlite*). In the <u>aligned</u> condition, the offset of the target word matched the intended word boundary (tãt#ʀøm), whereas in the <u>misaligned</u> condition, it did not (tã#tʀøm). In addition to the 72 experimental quadruplets, 72 similar CVCCV and CVCCVC strings such that their initial CVC portion did not form a word were created as fillers.

[1] From now, these are referred to as CVC for ease of description.

Table 1. Mean segmental durations (ms) as a function of the word-boundary location (#) and the lexical status of the CCV(C) final portion. Each cell (except C_{1b}) represents 72 data points. * indicates a reliable segmental lengthening due to the word-boundary location; C_{1b} refers to the 2nd consonant (glide or liquid) of the onset cluster in the 32 CCVC targets.

Conditions	C_1	C_{1b}	V_1	OB	LI	V_2	C_4
tãt#ʀu	91	29	135*	56	88*	99	
tã#tʀu	93	28	120	55	73	102	
tãt#ʀøm	87	28	140*	53	92*	122	89
tã#tʀøm	89	28	123	59*	77	119	87

All the phrases were produced by a female native speaker of French, unaware of the manipulation. She pronounced the phrases without separating the two words, that is with an *enchaînement*. The phrases were digitized (44.1kHz/ 16-bit), and the carrier strings were extracted using SoundDesigner.

Segment durational analyses were performed on the carriers (Table 1). Quite consistently with our previous findings in OBLI clusters [7], when the obstruent ended the first word (C#C), a substantial and reliable lengthening of the pre-boundary vowel (13.2%) and of the liquid (20%) was observed, independently of the lexical status of the second component of the phrase. Moreover, in the case of word-pseudoword phrases only, a reliable onset lengthening of the obstruent (11.3%) was found.

2.1.3. Procedure

Participants were required to press a button as quickly as possible whenever they heard a CVC word embedded initially in any string, and to repeat the word aloud immediately after. Carriers and fillers were mixed randomly in four experimental lists, Lexical status of the CCV(C) final portion and Alignment being counter-balanced across lists. Target structure was blocked.

2.2. Results and discussion

Two items were discarded from the reaction time (RT) and error (Er) analyses because they were missed by more than 50% of the participants. Chronometrical analyses were based upon RTs, measured from target offset, for correct responses only. Overall, target words were on average spotted 73 ms slower when their offset was misaligned (930 ms) with the intended word

Table 2. Mean reaction time (RT) measured from target offset and error rate (Er) in Experiment I (** indicates that the effect is statistically significant by participants and items).

Word CCV		Carrier type	RTs (ms)	Ers (%)
Target	Aligned	tɑ̃t#ʀu	818	11.4
	Misaligned	tɑ̃#tʀu	927	11.6
	Difference		109**	0.2
Nonword CCVC				
Target	Aligned	tɑ̃t#ʀøm	895	13.6
	Misaligned	tɑ̃#tʀøm	934	11.7
	Difference		39	-1.9

Table 3. Mean reaction time (RT) measured from target offset and error rate (Er) in Experiment II (** indicates that the effect is statistically significant by participants and items).

Word CCV		Carrier type	RTs (ms)	Ers (%)
Target	Aligned	tɑ̃#tʀu	1012	12.8
	Misaligned	tɑ̃t#ʀu	1037	12.9
	Difference		25	0.1
Nonword CCVC				
Target	Aligned	tɑ̃#tʀøm	1083	14
	Misaligned	tɑ̃t#ʀøm	1161	14.8
	Difference		78**	0.8

boundary than when it was aligned with it (857 ms). As shown by ANOVAs including Lexical status and Alignment as within-subject/item factors, this alignment effect was highly significant ($F_1(1,47) = 25.19$, $p< .001$; $F_2(1,69) = 11.02$, $p< .01$). More interestingly (see Table 2), the emergence of such an offset alignment effect depended upon the presence of a subsequent target competitor, as reflected by the interaction between Lexical status and Alignment ($F_1(1,47) = 4.04$, $p = .05$; $F_2(1,69) = 6.02$, $p< .02$). In the competition situation, a strong and highly reliable alignment effect was obtained: spotting the words was 109 ms slower when target offset was misaligned with the intended word boundary than when it was aligned with it ($F_1(1,47) = 26.95$, $p< .001$; $F_2(1,69) = 19.44$, $p< .001$). By contrast, in the absence of a target competitor, although the spotting latencies were somewhat longer (39 ms) in the misaligned condition, no reliable effect of alignment was found ($F_1(1,47) = 2.41$, $p< .13$; $F_2 < 1$). As regards accuracy, error rates were very similar across conditions, and no reliable effect or interaction were obtained.

The finding of a processing cost due to offset misalignment in the competition situation confirms the perceptual relevance of the segment durational variations (or of some phonetic correlates of them) observed within the materials, and thereby replicates Dumay et al.'s [7] results. Of more interest, the fact that such a misalignment cost is reduced to an unsignificant trend in the absence of target competitor clearly lends support to the idea that this effect results from the inhibitory competition that takes place between the CVC target and the CCV word activated at the following onset, as predicted by the SOSH hypothesis.

3. EXPERIMENT II

Experiment II used the same experimental stimuli to examine word-spotting performance for the CV (and CCV) initially embedded words. Thus, the aligned and misaligned conditions were reversed. Here, the lexical activation induced by the final portion of the carriers does not directly compete with the target, and the target is never disrupted by the intended boundary. However, as shown in Experiment I, when aligned, the lexical status of the CCV final portion affects the activation of the CVC longer initially embedded word. Hence, if the latter acts as a competitor for the CV target word, one would predict the same interaction of lexical status and alignment as in Experiment I, with an enhanced misalignment effect when the final portion is a word.

3.1. Method

3.1.1. Participants, materials and procedure

Forty-eight students participated in the experiment. All were native speakers of French, and reported no hearing or speech disorders. The target materials were those of Experiment I. As participants were required to spot the CV initially embedded words, other strings such that their CV initial portion never corresponded to a word were used as fillers. Otherwise, materials and procedure were identical to Experiment I.

3.2. Results

Six items missed by more than 50% of the participants were discarded from the RT and Er analyses. As in Experiment I, the main Alignment effect was significant on RTs: target words were spotted 52 ms slower when they were misaligned (1099 ms) with the intended word boundary than when they were aligned (1047 ms) ($F_1(1,47) = 7.91$, $p< .01$; $F_2(1,65) = 6.99$, $p< .02$). However, by contrast to Experiment I, no significant interaction between Alignment and Lexical status was obtained ($F_1(1,47) = 2.01$, $p< .2$; $F_2(1,65) = 1.84$, $p< .2$). The alignment effect showed nevertheless large variations across conditions, but quite surprisingly, they were in the opposite direction to those predicted, the larger difference being observed in the case of a nonword final portion. Despite the absence of an interaction, planned comparisons were carried out and revealed that the alignment effect was not significant when the target was followed by a word (25 ms; both $Fs < 1$), while it was highly significant when target was followed by a nonword (78 ms; $F_1(1,47) = 11.00$, $p< .002$; $F_2(1,65) = 6.18$, $p< .02$). Error rates were again very similar across conditions, and no reliable effect or interaction were obtained.

Joint analyses of Experiment I and II confirmed the clear difference in the patterns of RT results between both experiments ($F_1(1,94) = 5.82$, $p< .02$; $F_2(1,71) = 7.61$, $p< .01$). They demonstrated the target-dependent, inhibitory vs supporting, influence of the carrier CCV(C) final portion as a word, through the interaction between Target structure and Lexical status of the final portion ($F_1(1,94) = 4.47$, $p < .05$; $F_2(1,71) = 3.94$, $p = .0509$). Finally, they confirmed the higher difficulty to spot CV than CVC words, as reflected by the longer RTs to the former (1073 vs 894 ms; $ps < .001$) in the absence of any speed-accuracy trade-off (13.6 vs 12.1%; $Fs <$ or ≈ 1).

3.3. Discussion

Clearly the results do not fit with our predictions. First, we obtained a significant main effect of alignment. Since this effect is present in the nonword condition, it cannot be explained in terms of lexical competition. However, two alternative explanations can be proposed. One is that allophonic variations, presumably in the pre-boundary vowel, may directly modulate the activation of lexical candidates as a function of the boundary location (see also [6]). The presence of a 39 ms alignment effect in the nonword condition of Experiment I, although non-significant, could also be due to such a mechanism.

The other explanation is based on the Possible Word Constraint, introduced by Norris et al. [10]. According to this view, any lexical candidate is deactivated when the speech portion between its edge and a likely word-boundary is not a possible word. In the present misalignment conditions, the likely word boundary is located before the liquid, so that recognition of the initial CV word would leave the obstruent consonant unaccounted for. So *tɑ̃* would have its activation reduced in *tɑ̃t#ʀu* as well as in *tɑ̃t#ʀøm*, because the consonant *t* between its right edge and the perceived word boundary is not a possible word in French. Thus in this particular case, the application of PWC involves knowledge of possible word structures in French, together with the detection of likely word boundaries as proposed in SOSH. So, the alignment effects in the present experiment do not contradict the SOSH hypothesis.

The second intriguing outcome of Experiment II is the trend to an interaction between lexical status and alignment, such that the alignment effect was stronger in the nonword condition. This trend could be explained by the presence of possibly stronger boundary cues in the nonword than in the word condition (see the onset-lengthening of the obstruent for nonwords; Table 1).

Alternatively, the reduction of the alignment effect in the word condition could reflect a more complex interplay between the information sources involved in PWC than currently envisaged. PWC involves two elements, a candidate word and a likely word boundary. In the authors' view [10], the occurrence of an impossible word inhibits the lexical candidate but leaves the likely word boundary unchanged. Here, we suggest, a revision of the parse may be induced due to the lexical viability of the linking of the unaccounted obstruent to the final portion of the input.

4. CONCLUSION

This study re-examined the effects of offset misalignment on word-spotting performance obtained in our previous research. SOSH predicts that offset misalignment effects should be smaller than onset misalignment effects, except when combined with lexical competition. The results of Experiment I support this view. While the results of Experiment II did not conform to our initial prediction, we argue that they can be interpreted in terms of the interplay of SOSH and PWC.

5. ACKNOWLEDGEMENTS

This research was supported by grants from the Communauté française de Belgique (A.R.C. 96/01-203) and from the Swiss FNRS (Project 1113-049698.96). Many thanks to Laetitia Desmet for her voice and Cécile Fougeron for fruitful discussions. Correspondence to Nicolas Dumay, Laboratoire de Psychologie Expérimentale, Université libre de Bruxelles, Avenue F.D. Roosevelt 50, CP-191, B-1050 Brussels, Belgium. e-mail: ndumay@ulb.ac.be

6. REFERENCES

[1] 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.

[2] Content, A., Dumay, N., & Frauenfelder, U.H. (in press). The role of syllable structure in lexical segmentation in French. *Proceedings of SWAP*, Nijmegen, The Netherlands.

[3] Frauenfelder, U.H., & Content, A. (1999). The role of the syllable in spoken word recognition: Access or segmentation? *Actes des II^èmes Journées d'Etudes Linguistiques,* (pp. 1-8), Nantes.

[4] Dumay, N., Banel, M.H., Frauenfelder, U.H., & Content, A. (1998). Le rôle de la syllabe: segmentation lexicale ou classification? *Actes des XXIIèmes Journées d'Etude sur la Parole*, (pp. 33-36), Martigny, Suisse.

[5] McQueen, J.M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language*, **39**, 21-46.

[6] Smith, R., & Hawkins, S. (in press). Allophonic influences on word-spotting experiments. *Proceedings of SWAP*, Nijmegen, The Netherlands.

[7] Dumay, N., Content, A., & Frauenfelder, U.H. (1999). Acoustic-phonetic cues to word boundary location: Evidence from word spotting. *Proceedings of ICPhS '99*, vol. **1**, pp. 281-284, San Francisco, CA.

[8] Vroomen, J., & de Gelder, B. (1995). Metrical segmentation and lexical inhibition in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **21**, 98-108.

[9] Norris, D., McQueen, J.M., & Cutler, A. (1995). Competition and segmentation in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 1209-1228.

[10] Norris, D., McQueen, J.M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology*, **34**, 191-243.