A Comparison of Theoretical and Human Syllabification

Jeremy Goslin

Department of Computer Science

Submitted to the University of Sheffield for the degree of Doctor of Philosophy

February 2002

Abstract

Phonology is replete with theories and principles governing the process of syllable segmentation. However, with a myriad of conflicting theories available, a major problem facing researchers wishing to apply the syllable unit in models of speech processing is in selecting which of these theories to use in their definition of the syllable. In this thesis, an attempt has been made to judge the merits of the various principles and theories of syllabification applied to French, a syllable-timed language, by comparing the theoretical predictions of these theories with the experimental findings of a series of psycholinguistic syllable segmentation and perception experiments.

A number of factors have been found to influence syllable segmentation. These include, the nature of intervocalic consonant clusters and singletons, with preference given to minimal legal syllable onsets, excepting OBLI clusters, which are tautosyllabic. The aperture of the vowel at the nucleus of a syllable, with an open vowel inducing closed syllables, and vice versa. Also, orthographic bias, which, at the onset of literacy can influence syllabification decisions in metalinguistic tasks. In addition, examination of the differences between the syllable onset and offset detection revealed that, if there is a legal syllable onset before the first consonant of an intervocalic consonant cluster/singleton, then there will be a high degree of ambisyllabicity for this segment.

These findings suggest that listeners are taking advantage of a number of separate cues when segmenting speech into syllables. However, even when factoring these effects into the segmentation responses of subjects, there is still considerable variability in subject segmentation responses. By implementing the factors found to influence syllabification thus far, it is possible to suggest a set of preference rules which can predict where the boundary is *likely* to be located, and also to predict in which situations the location of the syllable boundary is likely to be most ambiguous.

Acknowledgements

I would like to thank everyone at the two laboratories where most of the research for this thesis was carried out.

For the formative years of this study, at the Speech and Hearing Research Group in the University of Sheffield, special thanks must go to my supervisor, Martin Cooke, without his continuous patience, support, and advice, this thesis would never have come about. Thanks also to Jon Barker, who showed me, and how right it was, that writing up is no bed of roses.

At the Laboratory of Experimental Psycholinguistics at the University of Geneva, I would like to thank the leaders of the FNRS project on word recognition, Uli Frauenfelder and Alain Content, for allowing a mere computer scientist to explore the world of psycholinguistics. Also, I would like to thank the attendant phoneticians of this group, Christine Meunier and Cecile Fougeron, for advice on phonetic matters, and for lending their eloquent voices in the production of the experimental stimuli. Thanks also to Liz and the rest of the group for putting up with my rapid English, and broken French.

Special thanks must go to Caroline Floccia, a constant guide and crutch throughout the PhD, without her constant advice, encouragement and insight the thesis would have gone awry long ago. Thanks also for the encouragement of my parents, and for my Dad to submitting to the unenviable task of proof reading, along with Caroline.

Finally, I must thank the hundreds of anonymous experimental subjects who agreed to venture into that subtle instrument of torture, the acoustic booth, to undergo the contemporary version of the inquisition in the name of science. Without these brave volunteers, willing to set their sanity aside for the greater good, this thesis would never have seen the light of day.

Contents

List of Figures

List of Tables

1 Introduction

Speech can be considered as a complex, an infinitely variable, stream of sounds resulting from the continuously-moving speech articulators. Since the utterances a person may be required to process are infinitely variable, the storage of direct representations of speech utterances would be impossible. Therefore, the analysis of speech is typically facilitated by breaking down the continuous process into a string of discrete units. A number of different candidates have been proposed to fulfil this role, including phonetic features, phonemes, and words. However, this study concentrates upon a single segmentation unit, the syllable, thought by some to hold a privileged position in the perception, production, and acquisition of speech.

1.1 The Syllable in Speech Perception

It is widely accepted that word recognition requires the search of an internal mental lexicon, which is generally considered to consist of up to 100,000 stored representations. Traditional, metalinguistic[1] psycholinguistic tasks, such as the lexical decision task, show that listeners take approximately 600ms to classify relatively frequent disyllabic words, a latency corresponding to the duration of the entire word. This finding might suggest that all of the acoustic information for a word is necessary to contact the lexicon and select the correct entry. However, in shadowing tasks, which are supposed to be capable of tapping onto on-line processing, in which listeners are asked to repeat a spoken target word as rapidly as possible, disyllabic words can be recognised after a delay of only 250 ms, that is, well before the end of the word (Marslen-Wilson, 1985). Similarly, gating tasks, in which longer and longer portions of words are presented until the target word is correctly identified, show that little acoustic information is required for the recognition of disyllabic words.

[1] Metalinguistic: aware of linguistic knowledge. A metalinguistic task allows the use of higher order knowledge and processing capabilities (at the lexical or post-lexical level of processing) to evaluate a linguistic task. These tasks differ from those considered as ‘on-line’ where responses are supposed to be based upon only pre-lexical, unconscious, processing (such as the fragment detection task).

Thus, it may be stated that the process of lexical access appears to be extremely rapid, especially given the large number of potential candidates stored in the mental lexicon.

This processing rate is even more astonishing when considering the *input variability* problem. No two productions of the same word are exactly the same, there are an infinite number of acoustically different tokens for a particular word, depending on factors such as speaker gender, speaking rate or accent, and the acoustic background, etc. Hence, the recognition of the spoken word not only requires a search amongst lexical candidates, but also the 'normalisation' of the input.

Perhaps the most widely accepted solution to this problem involves some kind of pre-lexical representation unit, built or retrieved from the sensory input and serving as the interface between the acoustic signal and the lexicon. Hence, word identification becomes a two step process. Firstly, the acoustic signal is normalised so that linguistic units may be computed and recognised. Secondly, the 'pre-lexical' code is used to search for the correct word in the lexicon.

1.1.1 Why a pre-lexical unit of processing?

According to the pre-lexical unit hypothesis, one or more sub-lexical representations mediate the mapping between the signal and the mental lexicon. This hypothesis, in the most part, is reliant upon the principle of economy. Instead of trying to match an infinite number of sensory inputs onto lexical entries, it is more economical to contact the lexicon with a finite set of abstract units retrieved from the speech input. However, one problem with this hypothesis is the question of how such representations can integrate prosodic, and other supra-segmental information. If an input representation consists of concatenated phonological units, such as a string of phonemes or syllables, it does not allow the encoding of supra-segmentation information, such as prosody.

Because of the deficiencies of pre-lexical segmentation units, a number of alternative models have been proposed which dispense with sub-lexical representations; information is extracted from sensory input and mapped directly onto the lexicon. However, even these models have to identify processing units. For example, according to the Cohort II model, a set of distinctive features are retrieved from the speech input, and used to contact lexical representations (Marslen-Wilson & Warren, 1994). Thus, the idea of building abstract representations before lexical access seems to be present in most word recognition models, although it varies in terms of the level of specification.

An additional debate surrounds the requirement for any form of *segmental* pre-lexical representation. Models proposed by Marslen-Wilson and colleagues state that it is not necessary to wait for the construction of a segmental unit before accessing the lexicon. Instead, it is suggested that the information used in lexical selection is continuously updated. If this were the case then features, rather than phonemes or syllables, would be the best candidate for the pre-lexical representation.

1.1.2 What kind of unit?

Many speech units have been proposed to serve as pre-lexical representations:

- temporally defined spectral templates (Klatt, 1980)
- distinctive features (Lahiri & Marslen-Wilson, 1991; Marslen-Wilson & Warren, 1994)
- phonemic segments (Marslen-Wilson & Welsh, 1978; Pisoni & Luce, 1987)
- demi-syllables (Samuel, 1989)
- morae (Cutler & Otake, 1994)
- prosodic units (Grosjean & Gee, 1987)
- the syllable (Mehler, 1981; Segui, Dupoux, & Mehler, 1990)

The relative suitability of each of these units for use in the pre-lexical level of representation can be expressed in terms of three criteria:

Immediacy of analysis: The smaller the unit, the more immediate, or faster, it can be mapped onto the lexicon. Proponents of the distinctive feature (Lahiri & Marslen-Wilson, 1991) have argued that this unit permits an almost continuous mapping of the signal onto the lexicon. However, too many lexical candidates reduce the initial economic advantage of postulating a pre-lexical level of representation. In contrast, larger units, like the syllable, can introduce long delays into the mapping process, especially if it is assumed that the lexical mapping only takes place when the end of the unit is reached.

Reliability of the analysis: The smaller the unit, the less information is present in the signal to support its identification. In addition, there is a relation between the unit's size and its variability. For example, the phoneme is subject to more variability than the syllable as there is less coarticulation between adjacent syllables than adjacent phonemes (Liberman & Studdert-Kennedy, 1978). Klatt (1980) has argued against the segmental unit, pointing out that segmental decisions can introduce errors of classification that misdirect the lexical access process.

Difficulty of the analysis: The complexity of speech analysis should, in principle, depend upon the unit inventory, which varies as a function of the size of the linguistic unit. The inventory size of different linguistic units varies considerably, by at least by a factor of 100. For example, the results of a lexical analysis of French (Goldman et al. 1996) found 30 phonemes as opposed to 3000 syllable types. If the recognition difficulty of a unit were measured as a function of the size of the inventory, then the unit with the minimum number of units would be the easiest to recognise. This assumes a constant function relating the relative difficulty in the discrimination of a particular unit, and the size of the inventory. However, using the example of the syllable, as inventory size increases across different languages, so does the diversity of the syllable structure (CV, CVC, etc.). Therefore, greater diversity in syllable structure could reduce the processing complexity due to an increase in the inventory.

1.1.3 How is the pre-lexical level of representation organised?

There are two complementary hypotheses concerning the organisation of the pre-lexical level of representation. The first suggests that the listener tries to match previously stored pre-lexical representations onto the incoming sensory input, just as the listener tries to match previously stored lexical representations to the sequence of pre-lexical representations (see Levelt & Wheeldon, 1994). The second hypothesis suggests that the listener computes appropriate pre-lexical units without being guided by stored representations.

In support of the first hypothesis, it is difficult to conceive of a recognition system that is not optimised to exploit the regularities of the speech input. Most syllables that are used in a given language are 'overlearned', both as articulatory gestures and as perceptual objects. As pointed out by Levelt and Wheeldon in a study of speech production (1994, pp. 246): "If these syllables scores are overlearned, it is only natural to suppose that they are accessible as such, that is, that they have a store of syllabic gestures for syllables that are regularly used in speech".

However, this hypothesis alone cannot explain why listeners are able to perceive units such as pseudo-syllables, that is, syllables that do not occur in a particular language, but are phonotactically plausible. If pre-lexical processing consisted only of matching incoming speech to stored representations, listeners would not be able to retrieve these units. Therefore, a "direct" route, just as in speech production or reading, must allow the listener to compute the syllable output from its atomic elements.

1.1.4 Experimental evidence

The experimental search for the existence of a pre-lexical *level* of processing cannot be easily dissociated from that of a pre-lexical *unit* of representation. Indeed, showing that listeners rely on a specified speech unit during pre-lexical processing is a privileged way of showing that there *is* a pre-lexical level of representation, and for specifying its architecture. Thus, most research on pre-lexical processing has aimed to determine the nature of the pre-lexical units.

Fragment detection paradigms

One heavily exploited paradigm for the investigation of pre-lexical processing is that of fragment detection (e.g. Bradley et al., 1993; Cutler et al., 1983,1986; Mehler et al., 1983; Norris & Segui, 1983; 1986; Mehler et al., 1981; Zwitserlood et al., 1993). In the original study, by Mehler et al. (1981), French listeners were asked to detect a number of pre-specified targets in bi-syllabic carrier words. Two types of target were used, fragments which consisted of the first syllable of the carrier word, and those that did not. For example, subjects were asked to detect the fragments /ba/[2] or /bal/ in the word /balance/ or /balcon/. Whilst /ba/ matches the first syllable of /balance/, /bal/ does not, with the reverse for the word /balcon/. Reaction times were found to be significantly faster when targets and words shared the first syllable than when they did not. This result, referred to as the ‘syllable effect’, was interpreted as providing evidence for the use of the syllable as a pre-lexical unit for access to the mental lexicon.

Following this seminal study, many experiments were performed using this paradigm. The majority of these studies concluded that the nature of the pre-lexical unit is determined, at last in part, by the phonological structure of the language, more precisely, by its rhythmic characteristics. For example, in syllable-timed languages, such as Spanish, a number of studies have successfully reproduced the syllable effect (Bradley et al.,1993; Sebastian et al., 1992). In similar studies, conducted with Japanese listeners, evidence was found to favour the use of the mora in pre-lexical processing (Cutler & Otake, 1994). The mora is a sub-syllabic unit traditionally used to describe Japanese rhythmic characteristics. Moreover, it has been shown that native listeners ‘export’ their segmentation strategy when processing words from another language. For instance, French listeners tested with English words displayed the same kind of syllabic effects than when tested with French material (Cutler et al., 1986). Therefore, the experimental evidence suggests that listeners develop a segmentation strategy related to their native language’s phonology, applying this strategy to any kind of incoming speech.

Although these results present a relatively coherent picture, they have not led to an accepted solution to the access unit debate. The first problem arises with the results of studies conducted with English-speaking subjects. In studies of English, listeners always failed to display the syllable effect (Cutler et al., 1986). One reason for this failure could be that, unlike French or Spanish, syllabic

[2] The IPA phonetic alphabet is used for all phonemic notation in this thesis.

boundaries are unclear in English, a phenomenon referred to as ambisyllabicity[3]. This hypothesis was tested by Zwitzerlood et al. (1993) using Dutch, a language that also exhibits ambisyllabic behaviour. They reported evidence for syllabification effects in Dutch listeners when presented with Dutch words having clear syllabic boundaries (such as "stremsel"), but also for words with unclear syllabic boundaries (such as "stre[m]ing"). These results suggest that the failure of English listeners to show the syllable effect cannot be linked solely to absence of clear syllabic boundaries in their native language. On the other hand, ambisyllabicity in English and Dutch differ, not only on the role of stress assignment, but also on the rules that are involved. In English, it is stress that is the main determinant of ambisyllabicity (Myers, 1987). In Dutch, it is dependent upon the quality of the preceding vowel, rather than on stress. Thus the disparity between the the behaviour of Dutch and English speaking subjects could be due to the differences between the languages at a more abstract level of phonological structure. Therefore the failure of English-speaking subjects to demonstrate the syllabic effect in fragment detection tasks could be explained by the specific ambisyllabicity rules in this language.

In addition, it would be false to suggest that the syllable effect is systematic through all syllable-timed languages. In the Sebastian-Galles et al.'s (1992) study, two Romance-language-speaking populations were examined, Catalan and Spanish adults. It was hypothesised that both populations should exhibit a syllabic effect in fragment detection tasks, but unexpected results lead the authors to offer the "acoustic transparency hypothesis". It was found that Spanish subjects exhibited a syllable effect in slower reaction times, of around 600ms, when reaction times were reduced to around 350ms the effect was no longer evident. However, Catalan subjects showed a syllable effect at 350 ms, but only when the to-be-detected syllable was accented (such as /cu/ in /cura/ or /cur/ in /cursi/). The authors of this study surmised that, depending on the acoustic transparency of the language spoken by subjects, and on the stimuli themselves, subjects can bypass the syllable in their analysis and detect the target fragment by relying on acoustic/phonetic cues. Catalan was said to be less transparent than Spanish, because of the presence of ambisyllabicity and the larger number of vowels in the former language. Thus, it is suggested that Spanish subjects would be more likely to bypass the syllable than those speaking Catalan. In addition, if the to-be-detected syllable is accented, its acoustic characteristics (length, amplitude) are more salient than those of a non-accented syllable. Therefore subjects should be more likely to bypass the syllabic computation if the target syllable is accented. Whilst these hypotheses may account for the experimental data, they do raise several additional questions. Firstly, the latency of the syllable effect in Spanish does not ensure the pre-lexical locus of the syllable effect. In addition, the model

[3] Ambisyllabic: a segment which belongs to two syllables

of syllabic processing suggested by the authors does not require that syllabic processing *must* be performed at the pre-lexical processing level, as claimed by the authors (claiming that the syllable will always be retrieved provided that enough time is given to the subject). Finally, this model would suggest that there are multiple routes open to pre-lexical and lexical processing.

Another important question, of direct relevance to all of these studies, concerns possible doubt as to the locus of the effects found in these experiments. That is, are the responses in these tasks computed pre-lexically, so as to mediate word recognition, or the result of a post-perceptual process involving intentional and conscious access to speech representations? As Kolinsky (1998) has pointed out, this task has not only been used for studying perceptual processes by the reaction times of subjects in detecting pre-specified targets, but also for exploring metaphonological development by measuring detection accuracy in infants or illiterate adults (Morais et al., 1989). It is assumed that the speeded version of the fragment detection paradigm is an on-line task (e.g. Mehler et al., 1981; Cutler et al., 1986), while the same procedure used as a measure of detection accuracy requires metaphonological judgements. However, at which detection time is it possible to decide whether the procedure taps into pre-lexical processing or post-lexical strategies? The results described earlier, obtained by Sebastian-Galles et al. (1992), could answer this question. It is suggested that syllabic effects emerging late (at around 620 ms) would be due to post-perceptual analyses of the stimuli, while the absence of such an effect with faster reaction times (around 350 ms) would reflect pre-lexical processing. However, this suggestion is of little help in interpreting the initial study of the syllable effect by Mehler et al. (1981), which found this effect in experiments with reaction times of around 360ms and 650ms. Therefore, according to the theory of Sebastian-Galles et al (1992), the syllable effect is found at both the pre-lexical *and* lexical levels of processing.

Perhaps, instead of criticising the pre-lexical hypothesis itself, it could be more pertinent to question the validity of the experimental task. In a study by Meunier, Frauenfelder, Content & Kearns (1997) the locus of the syllable effect was examined through the use of non-word stimuli with French speaking subjects. If the syllable effect were pre-lexical then the effect should be observed with non-word, as well as lexical stimuli. However, if, on the contrary, the effect taps onto post-lexical representations or strategies, the effect should not be observed with non-word stimuli. This study revealed that there was no syllable effect detected with non-word stimuli, despite great methodological care in controlling response latencies or stimuli properties. This outcome raises some doubts about the pertinence of fragment detection tasks in the examination of pre-lexical processing. However, as will be seen in the proceeding sections, there are other paradigms that are more pertinent to the exploration of this level of processing and that the results obtained so far seem to encourage the “psychological reality” of the syllable as a pre-lexical unit.

Other paradigms: phonological priming and speech migration

Problems surrounding the locus of the syllable effect, and the suitability of the fragment detection task in tapping into the pre-lexical level of processing require that other means be found for the investigation of pre-lexical speech processing. In this section three tasks are presented, each designed for the exploration of unconscious levels of speech processing.

An experimental paradigm recently used in the investigation of pre-lexical units is that of phonological priming. In auditory experiments, priming consists of presenting subjects with auditory (or visual) primes, followed by an auditory target, upon which a psycholinguistic task is performed (lexical decision, target detection, shadowing, etc.). Primes and targets are either phonologically (or semantically) related or not. The assumption is that the processing of prime-related targets should be easier than those that are unrelated, as the target is pre-activated by the prime, or that a mechanism involved in the processing of a shared property may still be activated in memory.

In a study by Radeau et al. (1995), French-speaking subjects were presented with phonologically related monosyllabic tri-phone primes and targets, sharing either their initial and final bi-phones. Subjects were asked to either produce a lexical decision, or a perform a shadowing task on the target. The results of these experiments showed that there was no priming effect in the initial overlap condition, but that there was a facilitatory effect in final overlap condition. The authors argue that this priming effect is due to the activation of the pre-lexical level of processing for three main reasons. Firstly, the size of the priming effect did not differ as a function of the lexical frequency of the item, suggesting that the effect was dependent upon pre-lexical processing. Secondly, the final overlap facilitation effect was modality specific, that is, it was only found when both the prime and target were presented aurally, not when using aural primes and visual targets (Radeau et al., 1994). Finally, the final overlap facilitation effect did not depend on the lexical status of the primes, since the size of the effect was similar when both words and pseudo-words were used as primes (Slowiaczek & Hamburger, 1992). Taken together, these results suggest that the phonological priming paradigm is of use in the investigation of pre-lexical processing.

However, before this paradigm could be used effectively, two issues have to the addressed. The first concerns the finding that initial phonological overlap does not lead to a priming effect. The second issue concerns the size of the unit(s) involved in pre-lexical processing. On the basis of the study by Radeau et al.(1995), it can be concluded that the priming effect was due to a rhyme overlap. The results do not tell whether a syllabic overlap would have led to stronger priming. This question has been addressed by the study of Dumay & Radeau (1997), who compared priming effects in rhyme and final syllable overlap in disyllabic French words. In this study a stronger priming effect was found with syllable overlap than rhyme overlapping. This finding is presented as evidence for the use of the syllable in the pre-lexical processing of French.

Pre-lexical processing has also been studied using an experimental paradigm based upon auditory illusions in the presentation of dichotic stimuli (Kolinsky & Morais, 1997; Kolinsky et al, 1995). In these experiments French speaking subjects were asked to detect or identify words created by blending two dichotic non-word stimuli. For example, subjects heard the non-word /biton/ in one ear and /cojou/ in the other. The results of these studies showed that word targets were reported more frequently when they were the result of syllabic fusion, rather than the fusion of other units. For example, the migration of syllables in the previous example frequently led to the perception of the two French words ‘bijou’ and ‘coton’. This paradigm was also used with Portuguese and Japanese listeners, producing a similar pattern of results to those found using the fragment detection paradigm. That is, that the size of the most frequently migrating unit was language specific. In Japanese, both the mora and the syllable appeared to elicit a high migration rate (Morais at al., 1996), whilst in Portuguese, the initial phoneme of words showed the highest migration rate (Kolinsky & Morais, 1993).

Finally, Pallier, Sebastian-Gallés, Felguera, Christophe, & Mehler (1993) adapted a phoneme detection task, first developed by Pitt and Samuel (1990), to explore whether it was possible to bias the attention of a subject towards the coda or onset of the first syllable of a word by varying the composition of stimuli lists. For example, inductors having a CVC as a first syllable, such as /ca*p*ture/, /su*r*face/ were alternated with test words made up of a CVC such as /segment/ or constructed using a CV such as /flagrant/ (the subject task was to detect the phoneme marked in italics). In this example the attention of the subject is biased to the coda of the first syllable. If subjects process words by using a syllabically based pre-lexical representation it is expected that subject’s responses would be faster on test words starting with a CVC than CV. This was the effect found with both Spanish and French speaking subjects. In addition, the same results were obtained with non-words, showing that the “syllabic” effect was not a by-product of lexical access, but rather that of ongoing pre-lexical processing.

In conclusion, it is suggested that the results obtained through the three tasks outlined in this section confirm the previous findings, that the syllable can be considered as the prime candidate for pre-lexical processing in Romance languages.

1.2 The Syllable in Speech Production

Similar to that of speech perception, traditional models of speech production postulate the existence of a sub-lexical level of processing (Levelt Model, e.g. Levelt, 1989). That is, following the stage of conceptualisation, there is a stage of grammatical encoding, which is followed by phonological encoding. The final two stages involve lexical access, where each word is supposed to be represented in two forms, the *lexeme,* used in grammatical encoding, and the *lemma,* used in phonological encoding. The lexeme contains semantic and syntactic information about the word, whilst the lemma contains phonological information. To suggest a broad analogy, the lexeme is *what* the word means, the lemma is *how* it is meant. The pre-lexical, or sub-lexical, unit would be activated at the phonological encoding stage. For example, in French, "it's my plane", becomes "c'est mon avion", which has the syllabic segmentation /se.mõ.na.vjõ/. The syllable /na/ does not belong to any word of the sentence, but was created at the phonological encoding stage along with the representation of the entire sentence.

An important source of evidence for investigations into phonological encoding, and the sub-lexical unit, comes from spontaneous production errors, such as "Irvine is quite clear", with clear being a mix of near and close. These errors allow an insight into which units are more likely to move, and hence into the nature of the units represented in the phonological stage. This suggests that the speech string is not computed as a continuous stream, but is the result of concatenated units smaller than the word itself. Evidence from these errors suggests that many units are likely to form the basis of the phonetic representation used during speech production, as errors in production occur with features, phonemes and syllables, with proportions varying from one language to the other. However, there is also evidence from studies using other experimental paradigms that support the use of the syllable as a privileged unit at this level of processing.

In a series of experiments with French speaking subjects, (Ferrand et al, 1996, 1997), target stimulus pairs were presented to subjects, who then had to pronounce or name the stimulus as rapidly as possible. It was found that responses were faster when the target corresponded to the first syllable of the stimulus (e.g. /ba/ followed by /balance/) rather then when the target did not match the first syllable of the stimulus (e.g. /ba/ followed by /balcon/). This, productional, syllable effect was found in both lexical and non-lexical stimulus items, and also when the stimuli were represented by drawings rather than orthographic, or oral, representations. These tasks were also repeated with English speaking subjects (Ferrand et al., 1997), who show little evidence in support of the syllable as a candidate in pre-lexical speech perception. In these experiments, two types of word were used, those having a clear syllable boundary (such as in ‘balcony’), and others with an ambisyllabic boundary, such as ‘balance’. Results showed a syllable effect for the first type of words, but not for words with an ambisyllabic boundary. However, it was also found that in the latter case, there were faster reaction times for both of the possible syllable boundaries (e.g. for both /ba/ and /bal/ in ‘balance’) compared to those where the target did not correspond to any possible syllable boundary in the word. It was suggested that these results supported the theory of ambisyllabic segments in English.

As in theories of speech perception, two of the most common hypotheses for the organisation of the sub-lexical level of speech production involve either the use of a mental syllabary, or suggest that a new representation is computed as required. Both of these hypotheses were examined in a study by Levelt & Wheeldon (1994). It was proposed that if there was a mental syllabary used in the sub-lexical level of speech production then there should be an effect of frequency within the syllabary. That is, the access of frequently used syllables should be faster than rare ones. Dutch subjects were asked to learn associations between abstract nonsense symbols and bi-syllabic words. Words were chosen with word and second syllable frequency crossover. The results of this experiment showed that production was faster for words with high frequency second syllables than those with rare second syllables. This effect was found to be independent of the frequency of the word itself, and the complexity of the second syllable. The authors suggested that these results were strongly in favour of the use of a mental syllabary in pre-lexical speech production.

1.3 The Syllable and Speech Acquisition

From the earliest age, infants can discriminate various acoustic, phonetic or prosodic changes within isolated syllables (e.g. Eimas et al, 1971; Jusczyk et al, 1992), whereas, for stimuli presented in a non-syllabic context, greater difficulties are encountered. For instance, there is no significant reaction in one-month-olds to the change from /pst/ to /tsp/, two stimuli that do not constitute proper syllables. However, they are able to distinguish the same contrast embedded in bi-syllablic utterances: /upstu/ vs. /utspu/ (Bertoncini & Mehler, 1981). In another example, newborns were successfully trained to discriminate the stimuli /paet/ and /taep/ in an operant conditioning paradigm, but could not achieve the same task with the stimuli /pst/ and /tsp/ (Moon, Bever & Fifer, 1992). More importantly, infants are also able to represent the syllable unit, or at least units organised around a vocalic nucleus, in multi-syllable words, as was shown by newborns' abilities to discriminate between lists of bi-syllabic and tri-syllabic words (Bijeljac-Babic et al., 1993). In this seminal study, using the standard High Amplitude Sucking procedure, newborns were presented with a list of phonetically varied CVCV French non-words. After habituation was reached, as attested by a decrease in High Amplitude sucking rates, a list of numerous phonetically varied

CVCVCV items were presented to half of the newborns (the experimental group), for the control group, the stimuli remained unchanged. A significant post-shift increase in HA sucking rates in the experimental group, when compared to the control group, indicated that newborns discriminated the list change. This response was interpreted, not only as showing discrimination between two lists of stimuli, but also categorisation. Subjects reacted to the change only because they could compare the two lists on the basis of a linguistic property, that is, the number of syllabic constituents (or the number of vocalic nodes). Given the large phonetic variability in both sets, it was extremely unlikely that newborns could detect a simple *item* change at the moment of stimulus shift. Before concluding in favour of an early sensitivity to the syllabic unit, the authors verified that newborns' reaction could not be attributed to a perception of duration change (since /CVCVCV/ is longer than /CVCV/), or to the perception of a change in the number of phonemes.

This study raised questions about the early sensitivity to other linguistic units such as the mora or the stress. Using a similar paradigm, Bertoncini et al. (1995) showed that newborns did not react to a change from bi-syllabic bi-moraic Japanese words (such as /ika/) to bi-syllabic tri-moraic words (such as /iNka/)[4], suggesting that they were more sensitive to the syllabic, rather than the moraic rhythm. Finally, van Ooijen et al. (1997) investigated newborns perception of the stress unit as it exists in English: a strong syllable, that could be followed by one of several weak syllables constituting a rhythmic unit in this language (for the moment ignoring the primary and secondary stress distinction). In their first experiment, newborns were presented with lists of monosyllabic stressed English words (such as 'foot') followed, after habituation, by a list of bi-syllabic weak-strong words (such as 'surprise'). The newborns displayed a significant reaction to the list change, showing that they were capable of perceiving the added syllable even though it was weak. In the second experiment, the authors investigated the possibility that newborns would consider the weak syllable to have the same rhythmic status as strong syllables. To this purpose infants were presented with a list of bi-syllabic weak-strong words (such as 'surprise') followed by a list of bi-syllabic strong-strong words (such as 'window'). In this study the infants did not react to the list change, indicating that they considered the stress pattern of words less pertinent than their syllabic organisation. Taken together, these results indicate that the syllable (or at least a vocalic node) appears as a privileged processing unit for early speech acquisition.

A viewpoint held by some phonologists is that the CV syllable forms the core syllable. That is, this syllable is used as a primitive and innate representational pattern used in speech production

[4] Japanese morae are subsyllabic units and can be geminate consonant (Q), a nasal sound (N), or a double vowel.

and perception. The CV syllable is then built upon, increasing in complexity during development, through the progressive branching of syllabic constituents (the coda will be defined, and then an extension in the number of available slots in the onset and coda position will take place). This was supported by analyses of Spanish and German infants production from the age of 0;9[5] to 2;1 (Lleo & Prinz, 1996). At the earliest stages, the majority of produced syllabic structures in both languages are CV. It is only afterwards that structures increase in complexity, following the order: CV-CVC-CVCC-CCVCC. However, even though 2-year-old children are already capable of producing correct 2-consonant cluster targets, mistakes are still made by the omission of one of the consonants. One finding of particular note is that the choice of the missing consonant will depend on the syllabic structure of the maternal language: from the age of 1;9, German children will omit the second consonant of the cluster more often than the first, with the reverse found for Spanish children. This effect could reflect that branching in German takes place from left to right, whilst in Spanish it takes place in the reverse direction. Thus, it appears that the syllabic representation is sensitive to language-specific constraints in production from the age of two. This, in turn, suggests that perceptual tuning to language-specific rules has taken place before this age. Another source of evidence comes from the demonstration that, from the age of 9 months, infants prefer to listen to lists of words that respect the phonotactic regularities of their maternal language than to words that do not (Friederici & Wessels, 1993). For instance, Dutch infants prefer to listen to words ending with a legal cluster in this language (such as *bref*) rather than those ending with an illegal cluster (such as *febr*). This suggests that infants have learned about the correct word boundaries at that age, and, as a consequence, they have an idea about correct syllabic structures.

At 5 years, it is often argued that the basic knowledge of syllabification rules is acquired (Fallows, 1981; Gillis & De Schutter, 1996). However, systematic discrepancies have been found between the performance of 5 year-old children and older children (around 8-10). These are usually accounted for by the unconscious application of spelling conventions to syllabification in spoken language, although it can also be argued that they are due to an increasing knowledge in the phonological components of the maternal language. For example, Gillis & De Schutter (1996) observed that 5-year-old children give ambisyllabic responses less often when the single intervocalic consonant of bi-syllabic words is a stop rather than a liquid, a nasal or a fricative. This is a result close to that found several times with English adults (e.g. Treiman & Danis, 1988). However, this factor no longer appears to play a role in the ambisyllabic responses of 8-year-old children. The authors offer two possible explanations for this modification: a radical change in childrens phonological representations of ambisyllabicity (lengthening vs. reduplication of the

[5] Age is represented in the form years;months; e.g. 1;4 = 1 year and 4 months.

consonant), leading to the equal treatment of all consonant types, or the application of spelling rules, since stops and other consonants are syllabified identically in written Dutch.

1.4 Defining the Syllable

As has been seen, the syllable has a prominent position in many models of speech processing. However, a question that has gone unmentioned thus far, but which is of central importance to the implementation of this segmentation unit, is, '*what exactly **is** a syllable?*'. In most of the studies where the syllable has been employed it remains undefined, on the assumption, it seems, that everyone knows what this highly intuitive unit is. However, everyone does *not* know, and the highly controversial subject of syllabic definition is usually overlooked.

Whilst the phoneme has the benefit of a relatively stable, clear and uncontroversial definition, being `the smallest distinctive unit within the structure of a given language´ (Trubetskoy, 1939), the definition of the syllable shares none of these benefits. Perhaps the most general, and least controversial, of the available definitions for the syllable is that of its common dictionary definition, that the syllable is '*a segment of speech that consists of a vowel, with or without one or more accompanying consonant sounds immediately preceding or following'*. Unfortunately, even this highly generalised pseudo-definition is not without its problems. This definition makes two claims, firstly that the vowel acts as the syllable nucleus, and secondly, that one or more consonants may be appended onto the syllable nucleus, those before known as the *onset*, after, the *coda.* Examining the first claim, there are complications that arise from the definition of the vowel, with either the phonetic or phonological distinction available. Taking the phonological distinction[6], vowels are defined as the set of syllabic vocoids (Pike, 1943). However, there is also a set of syllabic contoids, which can also function as syllabic nuclei. Modern phonological theory attempts to cut through these complications by simply assigning the binary feature [syllabic] to each segment, only [+syllabic] segments can be syllable nuclei. The second claim is less controversial, that a syllable may have an onset and coda, may be safely accepted at face value mainly because it specifies so little. Perhaps of greater interest is what remains unspecified.

Given the task of counting the number of syllables in an utterance in their maternal language naive listeners will have little difficulty, and will generally be in agreement. This task can be said to be roughly analogous to the counting of syllable nuclei. However, when listeners are

[6] As this study is largely concerned with the study of *phonological* theories of syllabification, phonological, rather than phonetic, definitions or terms will be used unless otherwise stated.

asked to state exactly where the syllable *boundaries* lie between those nuclei, then great difficulties are encountered with differences of opinion arising between listeners. It is this specification of syllable structure, or boundary location, that goes unstated in the original definition of the syllable that is most controversial.

Many theories have been put forward concerning syllable structure and syllabic segmentation. However, thus far, no consensus of opinion has been formed such that an accepted view of syllabic segmentation is available. However, an accurate prediction of where syllables begin and end is crucial to the definition of the syllable and its application in all fields of speech research. An inaccurate definition of the syllable could lead to experimental artefacts, especially in experimental paradigms for the investigation of the pre-lexical status of this segmentation unit. These experiments generally rely upon the location of syllable boundaries. If these boundaries are referenced incorrectly when selecting experimental stimuli, then it is natural to expect that any syllabic effects might be missed.

1.5 Research Aims and Thesis Overview

Phonological theory is replete with theories and principles governing the process of syllable segmentation. However, with a myriad of conflicting theories available, a major problem facing researchers wishing to apply the syllable unit in models of speech processing is in selecting which of these theories to use in their definition of the syllable. Experimental research can help to shed light upon this problem by asking the question: of the many principles and theories of syllabification proposed by phonologists, which ones do listeners follow? Although, as pointed by Lahiri and Marslen-Wilson (1991, p. 254), whilst it may be useful to postulate a direct mapping between phonological theory and a description of an individuals mental functions, "*Phonological theory cannot be interpreted literally, as direct description of mental representations, but it can be interpreted as specifying the functional properties of these representations*".

This thesis examines some of the most prominent phonological theories of syllable segmentation and compares their predictions with the segmentation responses of naïve listeners in a series of syllable perception experiments. In this way the merits of the various, conflicting, principles and theories of syllabification are examined to discover which provides the best representation of syllable segmentation in the perception of spoken French.

Chapter 2 reviews the most prominent phonological theories of syllable segmentation, including generic principles, and a number of algorithms designed specifically for the French language.

Chapter 3 presents a comparison of French syllabification algorithms, in practice, using lexical statistics, and against the segmentation responses of naïve listeners using syllable perception experiments. An examination of the consistency of syllable segmentation decisions made by experimental subjects is also discussed, both in general, and in relation to the segmentation differences between syllable onset and offset detection. In addition, the possible segmentation influences of orthographic bias and morphology are examined by comparing the experimental segmentation responses of pre-literate and literate subjects.

Chapter 4 reviews phonological theory concerning the possible influences of the syllable nucleus upon syllabic segmentation. Specifically, it examines phonological theories that concern the possible relation between syllable structure and vowel length, in English, or vowel aperture, in French. The *vowel aperture segmentation hypothesis* is presented for syllable segmentation in French. This hypothesis suggests that if the syllable nucleus consists of an open vowel this will induce the closure of the syllable, and vice versa.

Chapter 5 describes an experimental examination of vowel aperture production and perception to examine the underlying assumptions of the vowel aperture segmentation hypothesis. Namely, that French speakers can both differentiate between, and produce, open and close vowels in a particular syllabic context.

Chapter 6 describes an experimental examination of the influence of vowel aperture upon syllable segmentation, as proposed by the vowel aperture segmentation hypothesis. Two experimental paradigms are used in this examination, a metalinguistic task, first used in Chapter 2, and the fragment detection paradigm.

Chapter 7 presents the conclusions of this study, reviewing the factors found to influence the process of syllable segmentation in French.

2 Syllabification Models

“Si la syllable n’existait pas, il faudrait l’inventer”, Pulgram (1970).

“If the syllable did not exist, we would have to invent it”.

This chapter reviews a number of the most prominent generic phonological theories of syllabic segmentation, examining the commonalities and contradictions of the various principles of syllable boundary placement. Taking these principles as a basis of syllabic theory, a number of language specific implementations will also be examined, in this instance, for the French language.

2.1 General Syllabification Models

Probably the best starting point for a review of the major phonological theories of syllabic structure is the original pseudo-definition of the syllable given in Section 1.4. Here the problems with an exclusive relationship between vowels and syllable nuclei were briefly touched upon. In obviating this problem many phonological theories (e.g. Kahn, 1976; Hooper, 1972) define the nucleus of a syllable using the [syllabic] binary feature (Chomsky & Halle, 1968). The class of [+syllabic] units, includes all syllabic vocoids, but also a number of syllabic contoids, although a rigorous definition of this feature has yet to be achieved. With a one to one relationship between [+syllabic] segments and syllables the next step in the definition of the syllable is the formulation of principles for setting the boundary between each [+syllabic] segment.

Perhaps one of the simplest, and least controversial of these principles is that of the Obligatory Onset Principle (Hooper, 1972), or Principle of Maximum Open Syllabicity (Pulgram, 1970; Malmberg, 1963), also stated as a preference in the Head Law (Vennemann, 1988), and inherent, if not stated, in other theories of syllable structure (Clements, 1990; Selkirk, 1982). This principle is based upon the suggestion that the open syllable (a syllable with no coda) has historic or primitive significance in that all languages have open syllables (some only have open syllables) whereas no languages have only closed syllables (syllables with codas). Therefore, taking the example of a singular intervocalic consonant, the syllable boundary must lie before the consonant ([V.CV]) so creating the preferred open syllable.

Another of the most widely accepted principles of syllabification is that of the Legality Principle (e.g. Pulgram, 1970; Hooper, 1972; Kahn, 1976; Vennemann, 1986), which states that syllable onsets and codas are restricted to those phonotactically possible at word-initial or word-final positions. Evidence for this principle arises from two observations. The first is that no word medial cluster has been found which cannot be analysed as a possible word-final followed by word-initial cluster. The second observation is that, in cases where the syllabification of medial clusters is unclear, naive speakers do not produce syllabifications that involve clusters not found at word margins. However, it has also been suggested that the two laws of this principle, the Law of Initials, that legal onsets are found word-initially, and the Law of Finals, that legal codas are found word-finally, do not share equal status. In a study by Vennemann (1988) it was suggested that there were a greater number of exceptions in the Law of Finals than in the Law of Initials. This finding is also embodied in the Principle of Irregular Coda (Pulgram, 1970) which states that if an intervocalic consonant cluster cannot be divided into legal onset and coda, then the 'illegality' must be borne by that of the coda. This viewpoint is taken a logical step further in the Maximum Onset Principle (Kahn, 1976) which states that the syllable boundary must be placed before the *maximum* allowable legal onset, irrelevant of the legality of the syllable coda. An exception to this principle was proposed for the syllabification of intervocalic consonant clusters which suggests a hierarchy for the choice of the initial onset segment, with obstruents the optimum choice, followed by liquids and nasals, and finally glides (Hooper, 1972).

An alternative approach, built around studies of the 'sonority scale' (e.g. de Saussure, 1916), rather than analysis of phonotactic regularities, as those just discussed, is known as the Sonority Cycle (Clements, 1990). According to this principle, segments are ranked along a 'sonority' scale such that the preferred syllable type shows a sonority profile which rises maximally toward the peak and falls minimally towards the end of the syllable. Segments are classified in sonority using [syllabic], [vocoid], [approximant], and [sonorant] binary feature categories such that vowels have the highest sonority, followed by glides, liquids, nasals, and finally obstruents (similar to the hierarchy suggested by Hooper, 1972, but with the ranking of Nasals and Liquids switched). The Core Syllabification Principle states that segments are built upon the syllable nucleus (both onset and coda) as long as they are of a lower sonority to that of the preceding segment. However, as in previous theories, the construction of the onset of a syllable has priority over that of the coda of the preceding syllable. This can be seen in Figure 1, an example of the Core Syllabification Principle, showing the rise and fall of sonority for the word ‘implode’. In this case, the onset priority gives the syllabification /im.plod/.

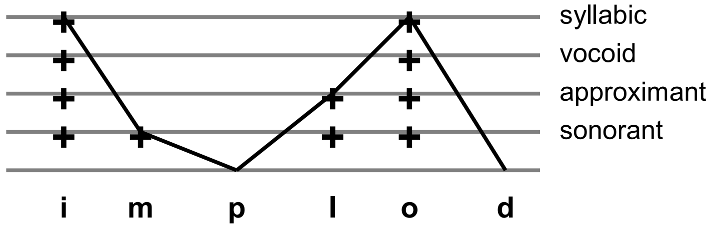


Figure 1 *Sonority Cycle for 'implode' (/implod/) showing phoneme sonority classifications and sonority curve*

Another, closely allied, but differently defined, theory is based upon a similar sonority scale, in this case inverted to form the Universal Consonantal Strength (Vennemann, 1988). This scale of sonority defines a larger number of consonanticity levels than the previous sonority scale, splitting, for instance, obstruents into plosives and fricatives (with the former of highest consonanticity), with descending values for high, mid, and low vowels. Syllable boundary placement is by means of the Contact Law and Contact Embedding preferences, which describe the same preferred sonority profile as that of the Sonority Cycle.

Thus far, all of the theories of syllable boundary placement have been based upon a flat, linear, representation of the syllable. However, an alternative suggestion (Selkirk, 1982) proposes a hierarchical structure of the syllable based upon the onset rhyme model, with the rhyme consisting of the peak (nucleus) and coda. Allowable syllabic form is dictated by overlaying a template of phonotactic constraints over this hierarchical structure specifying all possible syllable types for a particular language. An example of such a template can be seen in Figure 2, which encodes the gross characteristics of the syllabic structure of English using major phonological class features.

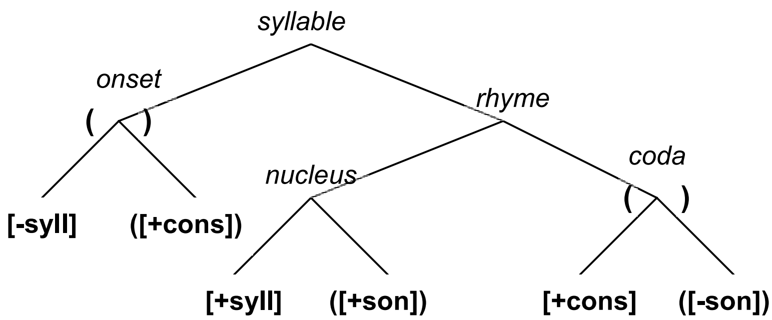


Figure 2 *Example syllable structure template for English, using onset-rhyme model, allowable syllabic form represented using phonological class features [syllabic], [consonantal], and [sonorant], branches in parentheses are optional (*from Selkirk, 1982)

Suggested advantages for such a representation include the ability to provide for explanations on phonotactic constraints without the need to resort to ad-hoc descriptions and the ease in which the description of the syllable may be fitted into higher order hierarchical organisations for the treatment of such phenomena as stress and rhythm.

One influence upon syllabification which has escaped attention in the description of the previous theories is that of stress. It is thought (Kahn, 1976; Selkirk, 1982; Vennemann, 1988) that the stress difference between a preceding, stressed, and proceeding, unstressed syllable, may affect the nature of the syllable boundary. One explanation of this phenomena suggests that, whilst in slow speech the types of syllabification rules discussed so far hold sway, during fast speech it is possible that medial consonantal segments found between stressed and unstressed syllables become *ambisyllabic* (Kahn, 1976), that is, belong to both the pro- and preceding syllables.

For example, and as shown in Figure 3, in slow speech the suggested syllabification of 'pony' follows the Obligatory Onset Principle, /po.ni/, however, under normal speech, because of the stress on the first syllable the segment /n/ becomes ambisyllabic, resulting in the two syllables /pon/ and /ni/. An alternative explanation for this behaviour is that of *resyllabification* (Selkirk, 1982) where, after initial syllabification, the syllable boundary may be moved as a function of stress, using the previous example this results in the change from /po.ni/ to /pon.i/.

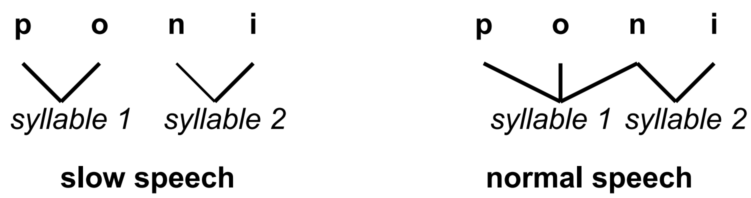


Figure 3 *Examples of the syllabification of 'pony' (/poni/) for slow and normal speech, with ambisyllabic segment for normal speech (*from Kahn, 1976*)*

This section has presented an overview of the major principles of syllabification available at the present time. Unfortunately, any theoretician hoping to bind these principles into a complete definition of the syllable is faced with a number of philosophical and practical conflicts between the different principles available. For example, the French word 'admettre' is syllabified /a.dmɛtr/ using the Core Syllabification Principle of the Sonority Cycle. However the Maximum Onset Principle places the syllable boundary at /ad.mɛtr/. In order to resolve such conflicts it is necessary to examine these different principles when applied to a specific language, such that an analysis may be made concerning the merits of the various conflicting principles of syllable boundary placement.

2.2 Syllabification in French

The French language has a number of advantages when it comes to research on the syllable. It is a syllable-timed language and is thought to have a relatively clear syllable structure, and therefore does not possess the problems surrounding that of stress-timed languages, such as English, regarding possible ambisyllabicity or resyllabification. The syllable also has special, psycholinguistic, relevance in French, with the discovery of the 'syllable effect' (e.g. Mehler et al., 1981; Cutler et al., 1986) suggesting the syllable as the primary perceptual unit used in lexical access, whilst similar studies in English have failed to find such an effect (e.g. Cutler et al., 1986).

Five different French syllabification algorithms are presented, chosen to reflect some of the differing general principles of syllabification outlined in the previous section. The first two (Coursil, 1992; Peereman, 1998) are based upon the concepts of sonority and consonanticity, each with differing hierarchies for the classification of French phonemes. The second pair (Dell, 1995; Laporte, 1993) suggests rules based upon analyses of French phonotactic regularities using the Legality Principle, both suggest special consideration for OBLI (OBstruent Liquid) clusters. The final selection is simply that of the Maximum Onset Principle, a complete model of syllabification in its own right, easily applicable to any specific language.

2.2.1 Coursil

Coursil's (1992) syllabification laws are inspired from that of the sonority scale, describing the same preferred sonority contour as that of the Sonority Cycle. However, following de Saussure's proposal, Coursil classifies segments into eight aperture categories, with a similar split of fricatives and plosives, and differing values for high, mid, and low vowels to that of the Universal Consonantal Strength. The classification of sonority across different segments can be seen in Table 1.

Sonority	Segments	Sonority	Segments
1	Minimal aperture: occlusives /p,t,k,b,d,g/	5	Glides
2	Fricatives	6	High Vowels /i,y,u/
3	Nasals	7	Medium Vowels
4	Liquids	8	Low Vowels /a,ã/

***Table 1** Sonority Levels of segments for Coursil Syllabification Algorithm*

Coursil (1992) states that syllable boundaries are assigned through a mechanism of double-plosion, using the definition of plosion originally used by de Saussure. Double-plosion occurs when an implosive segment is followed by an explosive segment. Binary values are assigned to each segment in the speech stream, 1 for an implosive segment, those whose aperture rank is higher than that of its following neighbour, 0 for the remaining segments. A syllable boundary is located every time this attribute changes from high to low (1 0). Because the start and end of utterances correspond to syllable boundaries, the start of every utterance is assigned a 0 (unless the first segment is a vowel, then it is assigned 1), the end 1. In the case of two consecutive phonemes of the same aperture rank, such as in geminate segments, they will have the values 1 0, and so mark a syllable boundary.

For example, the word 'moustique' (/mustik/) will be syllabified the following way:

	m	u	s	t	i	k
Aperture ranks:	3	6	2	1	6	1
Plosion values:	0	1	1	0	1	1

The syllable boundary is located between the string 10, that is /mus.tik/.

This mechanism operates similarly to that of the Core Syllabification Principle, describing the same preferred syllable boundary. However, some of the problems of *marked* demisyllables (Clements, 1990), containing violations of the sonority cycle in the form of sonority plateaus, are avoided using Coursil's syllabification principle based upon double-plosion.

2.2.2 Peereman

Peereman's (1998) approach is based upon the concept of consonanticity, similar to that of the Universal Consonantal Strength. However, Peereman suggests that the syllable is based upon semi-cycles, a semi-cycle of consonanticity followed by that of sonority, with the syllable boundary placed just before the consonanticity onset, as can be seen on the example shown in Figure 4. Segments are assigned consonanticity values according to the ten level scale seen in Table 2.

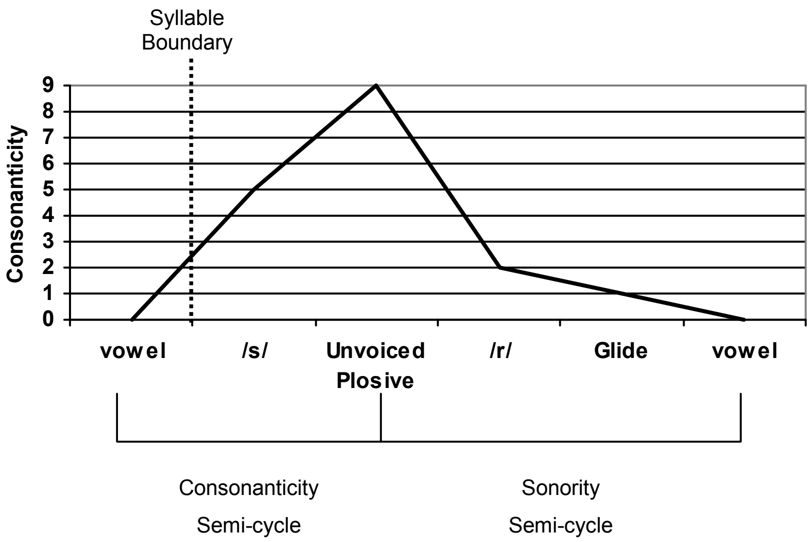


Figure 4 *Example of Peereman's consonanticity and sonority cycles consonan*

Consonanticity	Segments	Consonanticity	Segments
0	Vowels	5	/s/
1	Glides	6	Voiced fricatives
2	/r/	7	Unvoiced fricatives
3	/l/	8	Voiced plosives
4	Nasals	9	Unvoiced plosives

Table 2 *Consonanticity scale for the Peereman syllabification algorithm*

The suggested advantage for such a system, over that of a pure sonority scale, is its ability to present a framework where the /s/ + OB + (LI) clusters are tautosyllabic[7] (these types of cluster commonly form the start of words) as is the case in Selkirk's (1982) suggested hierarchical phonological template of the syllable for English. Peereman's analysis of consonantal attack, that is, sequences in which a segment has higher consonanticity than that proceeding, suggest that syllable onsets begin with segments which have a consonanticity value of at least 3 greater than that of the next. Using this assumption, plus the special treatment of /s/ segments outlined above, Peereman suggests the following two syllabification rules:

(i) If a consonant is at least 3 points higher than that of the following consonant they are linked to form an attack (syllable boundary)

[7] Tautosyllabic: belonging to a single syllable

(ii) An /s/ can be included in the attack if followed by a consonant of 3 points or higher (plosives).

As with the system proposed by Coursil, Peereman also solves the gemination problem, where segments form a sonority plateau. However, using the consonanticity scale proposed there are a number of cases in which no 'attack' takes place. Because the consonanticity levels of the segment /r/ and that of glides are below three (the minimum level of consonanticity level for an attack), then no attack occurs when these segments are placed in the context of singular intervocalic consonants (giving consonanticity levels of 020, or 010) or in Liquid Glide (LG) double intervocalic consonants (0310 or 0210). To avoid this problem an additional rule is proposed, that of the Obligatory Onset Principle, in these cases the syllabification will be /V.CV/ or /VL.GV/.

Using the syllabification example of the word 'moustique' it is possible to see the different treatment of /s/ clusters from that of Coursil.

	m u s t i k
Consonanticity ranks:	3 0 4 8 0 8

Here the syllable boundary is placed /mu.stik/ instead of /mus.tik/ using Coursil's system.

2.2.3 Dell

Dell's (1995) theories of syllabification stem from an analysis of the distribution of consonant clusters in French corpora, leading him to formulate the following rules for syllable segmentation:

(i) In an OBLI cluster the two consonants are tautosyllabic.

(ii) A postvocalic consonant is tautosyllabic with the preceding vowel, provided no conflict arises with (i) and provided that a prevocalic consonant remains tautosyllabic with the following vowel.

(iii) The coda contains at most one consonant.

The first of these rules gives special status to OBLI clusters, as in Kahn (1976) such that the syllable boundary is always placed before these clusters. However, Dell only uses a subset of possible OBLI clusters, excepting sequences in which both segments are coronal and whose first segment is a coronal fricative. This results in eight allowable obstruents, [p t k b d g f v], in combination with liquid segments, excepting /tl/ and /dl/. The second rule can be described as a synthesis of an analysis by means of the Legality Principle with that of the Obligatory Onset

Principle. The third rule ensures that tri-syllabic segments such as 'obstiné' are syllabified with the /st/ tautosyllabic.

2.2.4 Laporte

Laporte's (1993) method of syllable boundary placement was developed from a lexical study of syllable pronounceability. That is, when segmenting a pronounceable utterance in a particular language, a syllabification algorithm must operate such that all of the resultant syllables of that utterance are capable of being pronounced in isolation. However, Laporte goes on to question the validity of pronounceability, as it is not a formal notion, but based upon an intuition of a specific language. Therefore, he suggests an approximation of it, by lexical analysis of a large dictionary and the Laws of Initials and Finals of the Legality Principle, that is, syllables should divide such that the onset can be found as the first cluster or a word, and the coda as the last cluster of a word. Although he recognises that whilst these assumptions are false, (e.g. [gje] is pronounceable in [di.vyl.gje] but no French word begins with [gje]), it is suggested that they are of use as a first step in the production of a formal model of syllable pronounceability. Therefore, a number of algorithms, based upon these assumptions, were formulated and tested on a large French dictionary. Changes were made between each successive algorithm so as to reduce the number of words that were syllabified with illegal onsets or codas. The resulting algorithm generated illegal onsets or codas for only 22 out of 529 medial clusters tested, of these 4 contain the sequence [gj], and the other 18 clusters appear in only 32 words. This final algorithm is shown below:

(i) If a medial consonant cluster contains [p t k b d g f v] followed by [l r] it is treated as a single inseparable symbol.

(ii) Divide the cluster before the last symbol which is not a glide.

As can be seen, Laporte's syllabification algorithm is similar to that of Dell's, using a similar subset of OBLI clusters, excepting that in this case /tl/ and /dl/ are considered as tautosyllabic. The second rule of Dell and Laporte, though worded differently, have the same implications, both in their use of the Obligatory Onset Principle, and their treatment of glides (although explicitly worded in Laporte, Dell chooses to define his rules by means of consonants, such that glides will always be tautosyllabic with the preceding consonant). It is only in Dell's last rule, that the coda contains at most one segment, that these two algorithms significantly diverge, and why it is of interest to include *both* of these similar algorithms in the review.

2.2.5 Maximum Onset Principle (MOP)

The Maximum Onset Principle, as defined in a previous section, is perhaps the easiest of the general syllabification principles in its implementation. Though analysis of a large enough lexicon, such as BDLEX (Calmès and Pérennou, 1998) (~23000 words), it should be possible to approximate all possible onsets of French, and so calculate the Maximum Onset for any intervocalic consonant cluster.

To highlight a few of the differences between the French syllabification algorithms the various syllabifications for the words 'obscène' and 'extase' are presented in Table 3. One of the clearest differences between those algorithms based upon sonority and phonotactic regularities can be seen with the syllabification of 'obscène'. Here the sonority trough, and corresponding consonanticity attack, are before /b/. However, /bs/ is not a 'legal' onset, and therefore the remaining algorithms place the syllable boundary after /b/. With the tri-consonant cluster in 'extase' there is greater scope for difference amongst the algorithms. Here the special treatment of the /s/ segment in Peereman can be seen, differing from the sonority contour of Coursil, this effect is also evident in the differences between the Dell and Laporte algorithms, /s/ is tautosyllabic in Dell as the coda cannot be of more than one segment.

Model	Syllabification	
Dell	/ɔb.sɛn/	/ɛk.staz/
Laporte	/ɔb.sɛn/	/ɛks.taz/
Coursil	/ɔ.bsɛn/	/ɛks.taz/
Peereman	/ɔ.bsɛn/	/ɛk.staz/
MOP	/ɔb.sɛn/	/ɛk.staz/

Table 3 *Syllabification for the words 'obscène' and 'extase'*

3 A Comparison of Syllabification Algorithms and Syllable Perception

3.1 Introduction

In making a comprehensive comparison of the performance of the five French algorithms of syllabification it is preferable to base such an analysis, not on the examination of few contrastive examples, which simply serve to highlight the differences between models (which are usually evident in rule differences, or require an uncomfortable number of examples for a comprehensive analysis), but on the *impact* of those differences on the syllabification of the words of French. To such an end a *quantitative* analysis of the differences in syllable boundary placement was performed between the algorithms on a comprehensive lexicon of French words.

The first step in this process was a practical examination of these models, through their application on all intervocalic consonant clusters and singletons found in the French lexicon BDLEX. After lexical analysis it was necessary to use another, empirical, form of analysis for the clarification of syllable boundary placement. In order to extract the syllable boundary placements used by naive listeners, a metalinguistic syllable repetition experiment was performed using a broad range of stimuli. Using the preferential segmentation of subject's responses, that is, the segmentation given by the majority of subject responses, it was possible to form a test set of consonant clusters and singletons allowing an objective measure of agreement between human syllabification and theoretical predictions to be calculated. This comparison was then used to ascertain the merits of each of the syllabification algorithms.

However, one of the problems with current phonological theories of syllabification is that they are only capable of explaining preferential segmentation, that is the segmentation given by the majority of subject responses. By ignoring the remaining subject responses these theories can only present a surface representation of the processes behind syllabic segmentation. For a deeper insight into the processes of syllabification it is necessary to examine the whole distribution of segmentation responses. The comparison of subjects’ preferential responses and phonological models of syllabification is only the first step in search for an accurate model of syllabic segmentation. Further examination of the differences in the distribution of segmentation responses are also required to uncover some of the subtler influences upon syllabic segmentation which are not explained by current phonological theory. By examination of differences in syllabification consistency found between syllable onset and offset detection, and different categories of consonant cluster, it is hoped that it will be possible to paint a more complete picture of syllabic segmentation than would be possible through analysis of preferential segmentation alone.

Additional factors, such as the use of orthographic or morphological information, will also be investigated by means of additional metalinguistic experiments. These utilise a different experimental procedure, with the 'tapping task' used to elicit syllable boundary placement over both word, and non-word, stimuli in two groups of subjects, pre-literate children, and literate adults. A comparison of the segmentation responses between stimuli and subject groups will allow the investigation of the possible influences of orthographic knowledge and morphology upon the processes of syllabic segmentation.

3.2 Lexical Analysis of Differing Syllabification Algorithms

3.2.1 Method

To achieve a measure of similarity between the target syllabification algorithms a search was made of the BDLEX French phonetised lexicon (~23000 words) for commonly found intervocalic consonant singletons and clusters. This analysis found 431 different consonant singleton/cluster types in BDLEX with a total of 37221 occurrences (tokens) making an average of 1.68 singletons/clusters per word. These types are distributed as 19 single intervocalic consonants (26963 tokens), 208 bi-consonant clusters (9179 tokens), 172 tri-consonant clusters (974 tokens), 28 four consonant clusters (98 tokens), and three five consonant clusters (seven tokens). As can be seen, in French, as with most languages, there is a direct relation between the complexity of the consonant cluster, and thus, the syllable, and the frequency of its occurrence in words.

Taking these consonant clusters and singletons, each were syllabified using the five target syllabification algorithms with a comparison made of the resulting syllable boundary placement. An additional source of information for syllable boundary placement is BDLEX itself, as syllable boundaries are marked as part of the phonetic transcription of the words. One major difference between the syllabification used in BDLEX and that produced by the target syllabification algorithms is the influence of morphology. Whilst the syllabification algorithms do not acknowledge the influence of morphology, BDLEX enforces agreement between syllabic and morphological boundaries. For example, BDLEX normally syllabifies the OBLI cluster /bl/ as /.bl/ (137 tokens), however when appearing in the entry 'sublingual' the syllable boundary is moved to coincide with the prefix boundary, /b.l/ (1 token). Because the comparative analysis is performed on an inventory of clusters, not words, morphology cannot be taken into account. Therefore in this analysis all syllabic boundaries found in BDLEX are taken to be equally valid, even if they were the result of the influence of morphology. For example, in the comparison of syllabification agreement by type, if one algorithm produced the segmentation /.bl/ and the other /b.l/ then both would be in agreement with BDLEX, as both these segmentations of /bl/ can be found in the lexicon. However, in the comparison of syllabification agreement by token, 137 tokens would be in agreement with the first algorithm, and only one for the second.

3.2.2 Results

3.2.2.1 <u>Lexical Comparison of Syllabification Algorithms</u>

The percentage of syllabification agreement between algorithms, and BDLEX, for clusters of between two and five consonants, both by type and token, can be seen in Table 4 and Table 5. The comparison of single intervocalic consonants has not been included as all algorithms, and BDLEX (apart from those at prefix or composite word boundaries, as discussed previously), agree that these boundaries should be placed according to the Obligatory Onset Principle (V.CV).

	Dell			Laporte			Peereman			Coursil			MOP		
# of Cons.	2	3	>3	2	3	>3	2	3	>3	2	3	>3	2	3	>3
BDLEX	99	71	32	99	90	77	86	73	41	83	74	45	90	86	80
Dell				99	67	22	81	63	32	78	71	35	86	71	51
Laporte							82	68	35	79	75	51	85	79	58
Peereman										86	81	70	81	73	41
Coursil													79	68	45

Table 4 *Percentage agreement between BDLEX syllabification markers and syllabification algorithms on intervocalic consonant clusters found in BDLEX (% of types).*

	Dell			Laporte			Peereman			Coursil			MOP		
# of Cons.	2	3	>3	2	3	>3	2	3	>3	2	3	>3	2	3	>3
BDLEX	98	83	13	98	95	92	82	47	14	92	64	16	84	72	32
Dell				99	82	9	82	43	12	93	64	14	84	78	80
Laporte							82	44	11	93	66	19	84	69	24
Peereman										85	75	88	89	61	14
Coursil													87	48	16

Table 5 *Percentage agreement between BDLEX syllabification markers and syllabification algorithms on intervocalic consonant clusters found in BDLEX (% of tokens).*

To test the overall significance of syllabification similarity over the five algorithms a cross-tabulation of syllabification was calculated between pairs of algorithms. Each table was produced using the syllabification results of each algorithm for each consonant cluster and singleton type, with data represented as the number of consonants found in the syllable onset. This resulted in ten tables (pair-wise comparison between each of the five algorithms) each with five rows and columns (from one consonant onset to five consonant onsets). A McNemar test was then made on each cross-tabulation revealing that, with one exception, the syllabifications produced by the algorithms were significantly different from each other, with $p<0.05$. The only exception was the comparison of the Coursil and Peereman algorithms, where $p>0.05$. Whilst the syllabification produced by these algorithms are clearly not identical, as can be seen from the results, they do not appear to be significantly different.

Examining the similarities between BDLEX and that of the algorithms it was found that the strategy used for the placement of syllable boundaries in this lexicon strongly resembles that of Laporte. Exceptions to this mainly involve the treatment of the /s/ segment. In Laporte the /s/ has no special status (as in Peereman, or inherent in Dell) so the cluster /psk/ is syllabified as /ps.k/. However, in BDLEX, this is syllabified /p.sk/. The treatment of these segments in BDLEX is not as consistent as that described in the syllabification algorithms, for example, /bsk/, is syllabified in agreement with Laporte, as /bs.k/. For clusters with more than three consonants this difference becomes more apparent as a greater proportion of these clusters use the /s/ segment. This effect is evident in the greater levels of agreement between BDLEX and MOP than Laporte for these clusters when analysed by type. However, these clusters are of relatively low frequency, as agreement by token between MOP and BDLEX is only at 32%, whilst for Laporte and BDLEX it is at 92%.

3.2.2.2 Syllabification Algorithm Agreement relative to Cluster Length

Comparing the syllable boundary placement offered by the five algorithms of syllabification, it can be seen that agreement levels remain relatively high for double intervocalic consonant clusters, where there are only two acceptable boundary locations. However, as the length of the consonant cluster increases, the agreement reduces. For two algorithms that share a similar background, those of Dell and Laporte, there exists almost total agreement in the syllabification of double intervocalic consonant clusters (excepting only /tl/ and /dl/ clusters). Agreement in longer clusters is greatly reduced, showing that relatively small and simple differences in syllabification rules can reap large differences when applied across a wide range of stimuli.

These findings are more easily appreciated if the consonant clusters are arranged as a factor of the maximum number of algorithms that are in agreement. For example, the cluster /ks/ is syllabified /k.s/ by Dell and Laporte, but /.ks/ by Peereman and the Maximum Onset Principle, therefore the algorithm agreement level for this cluster is two. As can be seen, in this study the Coursil algorithm has been removed from the analysis, as Peereman and Coursil are so similar only one of these algorithms have been included in the confidence calculations. Because all single intervocalic consonants are syllabified the same way by all algorithms there is but a single data point, 100% of singletons with a algorithm agreement level of 4. Table 6 represents the distribution of agreement amongst syllabification algorithms, showing the percentage of consonant clusters (by token and type) for different levels of algorithm agreement.

Number of Algorithms in Agreement	Number of Consonants in Cluster			
	1	2	3	>3
1	0 , 0	0 , 0	0 , 0	6.5 , 2.1
2	0 , 0	7.2 ,11.3	18.0 , 28.9	48.4 , 78.9
3	0 , 0	18.3 , 10.1	30.8 , 31.0	29.0 , 12.6
4	100 , 100	74.5, 78.6	51.2 , 40.1	16.1 , 6.4

Table 6 *Percentage of cluster types & tokens (types, tokens) covered by different levels of algorithm agreement for different lengths of consonant cluster*

As can be seen, Table 6 shows that as the length of the consonant cluster increases there is less agreement between syllabification algorithms. Separating consonant clusters in this way may have an additional benefit, in that the level of agreement amongst syllabification algorithms could be used as a measure of syllabification *confidence.* Take, for example, the cluster /lm/, this is syllabified as /l.m/ by all of the algorithms, and so has an agreement level of four. For this cluster the different algorithms of syllabification converge: there is no algorithm that suggests that this cluster should be syllabified in any other way. However, taking the previous example, /ks/, considerable controversy was found amongst the syllabification algorithms, two predicting the syllabification /.ks/ and the remaining two /k.s/. Because, at the present time, there are no means of deciding which of the algorithms is 'correct', there is no way to decide which of these two segmentations should be used.

One possible solution would be to use the syllabification given by the majority of algorithms. However, this assumption requires that the segmentation differences between algorithms are equally distributed. Given the wide theoretical differences between some algorithms, and similarities between others, the validity of this assumption is highly unlikely. A workable alternative is the measure of confidence. For those clusters, like /lm/ where there is no controversy surrounding syllable boundary placement it is possible to have a high level of confidence in its syllabification. For others, like /ks/ there is less confidence in the segmentation as the level of agreement is lower. The use of confidence, while not resulting in a definite syllable boundary placement, is useful in highlighting the types of consonant cluster that present the greatest problems and so require further study. This measure could be useful in applications where the correct assignment of the syllable boundary is of utmost importance. In this case, if there is leeway in the choice of consonant cluster under analysis, then priority should be given to those in which we have the greatest syllabification confidence. It is interesting to note that, taking all intervocalic consonant clusters and singletons, just over 61% of types and, more importantly, 91% of tokens, are syllabified with all algorithms in agreement, and so at the highest level of confidence. This finding also suggests that those consonant clusters that give greatest problems to syllabification are significantly less frequent in the lexicon than their less controversial counterparts, although it must be stated that this is mostly due to the frequent occurrence of single intervocalic consonants.

3.2.2.3 Agreement with the Legality Principle

Thus far the performance of the five algorithms of syllabification have been compared. It may also be useful to analyse the status of the syllable onsets and codas that are generated using these algorithms regarding their agreement with the Legality Principle. A measure of legality was produced by searching BDLEX for all possible word initial and final consonant clusters/singletons. This resulted in 128 possible onsets, and 109 codas. Analysis of legality agreement for the five algorithms was produced by syllabifying all possible intervocalic consonant clusters and singletons and comparing the resultant syllable onsets and codas with those from the previous analysis. The percentage of those resulting in legal onsets, codas, and both onset and coda can be seen in Table 7.

Again, single intervocalic consonants were not included, as the syllabification /V.CV/ produces legal onsets and codas for all consonants.

	Number of Consonants in Cluster								
Model	2			3			>3		
	Onset	Coda	Both	Onset	Coda	Both	Onset	Coda	Both
Dell	98	99	98	76	98	76	52	100	52
Laporte	97	99	97	96	96	93	100	84	84
Peereman	89	99	89	74	98	73	45	97	42
Coursil	84	99	83	75	98	74	55	97	52
MOP	100	99	100	100	99	99	100	100	100

Table 7 *Percentage agreement levels between syllabification algorithms and Legality Principle for Law of Initials (Onset), Law of Finals (Coda) and both Laws for different lengths of consonant cluster.*

One of the more obvious results of this analysis is the disparity in Legality agreement between onset and coda, averaging across all algorithms and consonant clusters, 87.9% of all generated onsets were legal, versus 98.5% of codas. Therefore it appears that the application of the algorithms is contrary to that suggested by Vennemann, who stated that the Law of Initials is stronger than that of the Law of Finals. As expected, those algorithms based upon sonority rather than phonotactic regularities have particularly low agreement with the Law of Initials. However, for consonant clusters with greater than two consonants, the Dell algorithm also has very low agreement, particularly when compared to that of Laporte that has higher Law of Initials agreement than that of the Law of Finals. This finding highlights the main difference between these algorithms, in that Dell only allows a single coda, this increases the complexity of onsets for longer consonant clusters, and so reduces the chances of matching with word initial clusters. The Maximum Onset Principle, inherently, has total agreement with the Law of Initials, but also has the highest agreement with the Law of Finals, and thus the highest total Legality agreement. This is particularly interesting as this algorithm pays no heed to the possible phonotactic requirements of the coda and yet obeys the Law of Finals to a greater degree than the other algorithms based upon analysis of allowable onsets *and* codas.

3.2.3 Conclusions

To summarise, it appears that most algorithms, even those based upon similar theoretical lines, with only minor rule alterations, will yield large differences in syllable boundary placement, especially as the consonant cluster increases in complexity. The exceptions being the two algorithms based upon the concepts of sonority, those of Peereman and Coursil, which produced statistically similar patterns of syllable boundary placement. These finding suggest that of the five syllabification algorithms tested those of Dell, Laporte, and the Maximum Onset Principle are significant in their own right. These algorithms are worthy of individual examination and testing as there were significant differences in their practical application. It could be said that, as the algorithms of Peereman and Coursil are not significantly dissimilar when applied to a large lexicon, they should be merged in some way, or one dropped in favour of another. However, whilst 83% of the intervocalic consonant clusters and singletons (and 96% of tokens) were syllabified in the same manner by these two algorithms, this still leaves 69 consonant clusters (1616 tokens) where the algorithms were in disagreement. If these clusters were rare, and so would have negligible impact on the syllabification, or could be traced to artefacts such as loan words, then it is possible that the disparity amongst these clusters could be overlooked. Unfortunately, this is not the case, therefore the same problem remains, although reduced, which syllabification decision do you select for the remaining consonant clusters?

Whilst the issues of syllabification similarity are acute with the Peereman and Coursil, there are also a number of clusters where the remaining algorithms also converge. In these cases the number of algorithms in agreement can be used as a measure of confidence, the larger the degree of convergence, the more confident we can be that that particular syllabification is correct. However, whilst this measure can be useful in the analysis of syllabification it cannot help in the definition of the syllable, only inform where the problems might lie. In all cases what is required is a form of benchmark, a set of consonant clusters in which we can be certain of the syllable boundary placements such that the algorithms may be tested against an impartial measure, instead of the subjective analysis produced when they are compared against each other. Therefore it is necessary to turn to the only natural source of information for syllable boundary placement, the naive listener, in order to produce such a benchmark.

3.3 Experiment 1 – Syllable Repetition in Adults

There has been growing interest in the use of metalinguistic tasks for the comparison of phonological theories of syllabification in order to establish which of the many principles of syllabification are followed by listeners. Thus far, a large majority of this work has been performed on the English language (Fallows, 1982; Treiman & Danis, 1988; Treiman & Zukowski, 1990; Treiman et al, 1992; Titone & Connine, 1997), and, to a lesser extent for the Dutch language (Gillis & DeSchutter, 1996). It is only recently that similar studies have been made for French (Content et al, in press).

In these studies a number of different tasks have been suggested to elicit syllable boundaries from experimental subjects, and may be split into two categories, orthographic, and oral. Orthographic tasks usually involve the presentation of the written form of word, or non-word, for which they have to select the correct hyphenation (Treiman & Danis, 1988; Treiman & Zukowski, 1990, Treiman et al, 1992). Oral syllabification tasks are many and varied, examples of such are first or second syllable doubling (Fallows, 1982) (such that the word 'cobra is repeated as 'cocobra' or 'cobrabra'), or the tapping task (Liberman et al 1974) where subjects repeat the word tapping at each syllable and inserting a pause at syllable boundaries. Two other tasks, syllable reversal (Treiman & Danis, 1988) ('cobra' repeated as 'bra [pause] co'), and first or second syllable repetition (Treiman et al, 1992; Content et al, in press), are very similar, excepting that the repetition of the first and second syllables may be blocked separately for the latter task.

Unfortunately there are a number of unwanted influences that may affect some, or all, of these tasks to a greater or lesser degree. These influences are due to orthography, contingency, and off-line processes. The first of these concerns the possible influence of orthographic division rules, learnt at school, on syllable boundary placements (Pulgram, 1970). Although this is considered of greater importance on written tasks, the orthographic form of a word used in oral tasks may also have an influence upon its syllabification (Treiman & Danis, 1988). The second contaminating influence is that of contingency, that is, both syllables in a segmentation task are extracted in close proximity. This influence, particularly prevalent in the syllable doubling, or tapping task, and to a lesser degree in the syllable reversal task, suppresses the possibility of an ambisyllabic response, as the close proximity of the two syllables makes such a response sound 'unnatural'. For example, when segmenting the word 'cobra' the subject may repeat 'co bra' or 'cob ra', but an ambisyllabic response, 'cob bra', is unlikely because the repetition of /b/ would not be natural to normal speech. The last influence is related to the metalinguistic nature of all of the previous oral syllabification tasks. As the syllabic representation used in real-time speech processing is of most interest, the role of metalinguistic knowledge, or strategies, in syllabification responses is a potential source of experimental contamination. For these reasons most psycholinguistics experiments that address the issues of segmentation work with either excised, or isolated, fragments of speech (usually words or non-words). By using these types of experimental stimuli aspects of continuous speech, such as supra-segmental prosodic, pragmatic, or syntactic information, that could mask, or contaminate segmentation decisions are minimised. The rationale behind this strategy is that if a pre-lexical segmentation mechanism can be shown to operate in an isolated speech context then, assuming that the pre-lexical speech processing is constant across all types of speech input, the same mechanisms would also hold for continuous speech.

Perhaps the optimal task, that which reduces these contaminating influences to the minimum, is that of syllable repetition. In this task non-word stimuli can be used to reduce the possible influence of orthographic knowledge, by separating the repetition of the first and second syllables into different experimental blocks the problem of contingency is virtually eliminated. Finally, by running a speeded task, giving the subjects only the minimum amount of time to repeat the required syllable, it is hoped that influence of metalinguistic knowledge or strategies will be kept to a minimum. However, it is acknowledged that this factor may only be reduced with the use of an on-line task.

3.3.1 Stimuli

3.3.1.1 Test items

To present the widest range of possible stimuli, suitable for a general study, and comparison, of syllable boundary placement, stimuli were selected from all consonantal robust feature classes (Nasal (N), Fricative (F), Liquid (L), and Plosive (P)), excepting glides, which were not used because of their association with the syllable nucleus. Using these robust feature classes, stimuli were organised into four singleton and 16 double (e.g. FN) (a full list of these categories can be seen in Figure 6), and 3 triple (PFP, LPL, and NPL) consonant categories. Depending upon the availability of consonant clusters, these categories contained between 1 and 3 different clusters. Each consonant cluster/singleton was repeated a number of times such that there were always 6 tokens per category. For example, the consonant category FF contains only one legal cluster, /sf/, whilst FP contains many (e.g. /ft/, /sp/), to make 6 tokens per category /sf/ must be repeated six times, but for FP three clusters were selected, each repeated twice to make 6 tokens. This arrangement increases stimulus variation, and experimental coverage, whilst the number of observations per consonant category remains constant.

To create a bi-syllabic non-word, a random selection of vowels was placed at the start and end of each consonant singleton/cluster (forming VCV, VCCV, and VCCCV stimuli) taken from [/u/, /i/, /a/, /y/]. These vowels were chosen because there do not appear to be any constraints on the distribution of these vowels due to syllabic structure. Whilst there are constraints on the placement of the mid vowels due to syllable structure (e.g. vowel harmony), more specifically, syllable aperture (closed syllables having a coda, open syllables with no coda) (Tranel, 1988), this does not appear to be the case for the remaining French vowels. Therefore, it is thought that these vowels are

[8] The IPA phonetic alphabet is used for all phonemic notation in this thesis.

less likely to have an effect on the syllabification behaviour of the subjects than the mid vowels, which have been directly linked to syllable structure. This organisation of stimuli resulted in 138 tokens, with 57 clusters/singletons, and 23 feature categories. Examples of such stimuli are /igla/ or /upy/. A full list of all stimuli can be seen in Appendix A.

3.3.1.2 Training items

A total of 10 training items were also generated in similar form to that of the test items using the same distribution of VCV, VCCV, and VCCCV stimuli used in the test set, but using different consonant cluster/singletons.

All stimuli were produced by a female adult monolingual, naïve (that is, they had no knowledge of the aims or procedure behind the experiment), native French speaker (Parisien accent) from a randomised list of non-words represented phonetically using the IPA alphabet, with a pause of three seconds inserted between the production of each stimulus. These non-words were recorded using a Sennheiser ME80 microphone and Rane MS1 microphone pre-amplifier onto Sony ZA5ES DAT recorder. Afterwards the stimuli were low-pass filtered at 11kHz and then digitised with a fidelity of 22kHz with 16 bits per sample.

3.3.2 Procedure

The experiment was performed automatically using the Psyscope experimental software running on a Macintosh computer with accompanying button box. Audio stimuli were reproduced using this software through an audiomedia sound card and presented to subjects using Beyer Dynamic DT100 headphones. During the experiment subjects were seated in an amplisilence acoustic booth.

Subjects were asked to repeat either the first or second part of the bi-syllabic stimuli. Experimental stimuli were arranged into 3 blocks, each of which was presented under two conditions, that is, the repetition of either the first syllable (Condition 1) or second syllable (Condition 2) of the stimuli. Stimuli order for each block was randomised for each presentation. At the start of each experimental block the subjects heard a double tone. This indicated that the subjects should look to the lights above the two buttons of the button box. These lights indicated the condition of the proceeding block (repeat first or second part of the stimuli). Subjects acknowledged the condition by pressing the relevant button. Stimuli in each block were presented on a continuous basis with a short warning tone heard 300ms before each stimulus, one every two seconds without pause until the end of the block, where subjects were invited to take a short pause. Subjects were asked to repeat either the first or second part of the stimuli, according to experimental condition, immediately upon presentation, with the quickest possible response. The experimental condition alternated on each successive block, with blocks ordered such that block repetitions (for the 1st and 2nd condition) were never adjacent to each other.

At the start of the experiment instructions were read to the subject from a pre-prepared information sheet and any questions answered. Two short training blocks of 10 stimuli were then presented, one for each experimental condition. After the training block further questions were invited before the subjects moved onto the main experimental blocks.

All subject responses were captured using a Sennheiser ME80 microphone and Rane MS1 microphone pre-amplifier onto the right channel of a Sony ZA5ES DAT recorder. The auditory stimuli presented to the subjects was also recorded on the left channel of the DAT.

At a later date the responses were transcribed by a trained phonetician using a list of the phonetic representations of the stimuli, and the subject responses and stimuli previously recorded onto DAT. The phonetician heard the stimuli over DT100 headphones, with the stimuli presented in the left ear, and the response in the right, as recorded onto the DAT. The phonetician marked a syllable boundary for each stimulus if the response coincided with the phonetic representation of the stimuli (that is, if there was no disparity between the phonemes used in the response and that of the corresponding fragment of the stimulus), or, if not, marked the response as erroneous.

3.3.3 **Participants**

All 22 participants (19 female, 3 male) were students of the Université de Genève, with an average age of 22;7 (ranging from 18 to 32), and were native speakers of French with no known hearing defects. They received course credits for their participation.

3.3.4 Results

3.3.4.1 Error rates

Errors consisted of missing responses, mispronounced repetitions (using phonemes not found in the original stimuli) and repetition of the total stimuli (both syllables). Error rates averaged at just fewer than 2% of the possible responses, there was no significant effect of consonant cluster on error rates, with the highest subject error rate at 8%.

3.3.4.2 <u>Preferential Segmentation</u>

Segmentation responses, shown in Figure 5 and Figure 6, were analysed separately for each of the experimental conditions (first and second syllable repetition). To test the significance of the preferential segmentation responses a χ^2 goodness of fit test was used, an appropriate measure for the analysis of categorical frequency data. The χ^2 analyses were conducted on the distribution of subject responses for each possible segmentation decision for each of the single, double, and triple consonant stimuli.

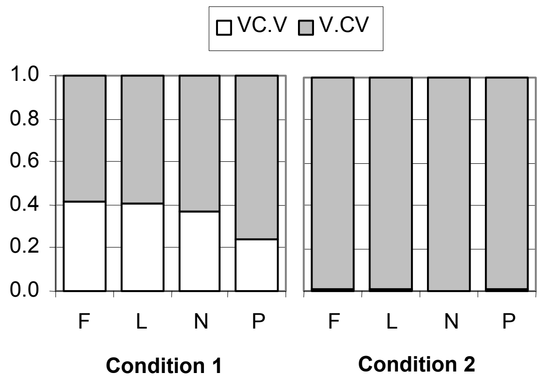


***Figure 5** Distribution of subject syllabification responses in Experiment 1 for conditions 1 and 2 for all robust feature classes of single consonant stimuli*

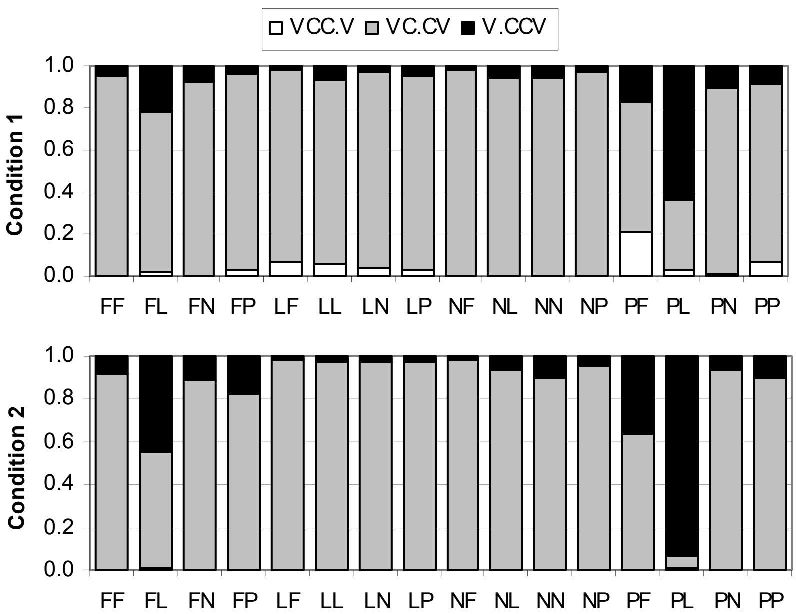


***Figure 6** Distribution of subject syllabification responses in Experiment 1 for conditions 1 and 2 for all robust feature classes of double consonant cluster stimuli*

For each stimulus a comparison was made between the numbers of subject responses for each possible segmentation decision (two possibilities for single consonant stimuli, /V.CV/ and /VC.V/, three for double consonants, and four for triple consonants) and the total number of responses, evenly distributed amongst the possible categories. Dependent upon the stimulus under analysis, as some clusters/singletons were presented to the subjects twice, others six times (see section 3.3.1.1), there were between 44 (two presentations for each of the 22 subjects) and (six presentations) 132 observations for each analysis. In this test the χ^2 measures are used to ascertain whether the preferential segmentation decisions for each stimulus is significantly different to a random a distribution of responses. In most cases, the χ^2 was higher than the critical level (at $p= 0.05$, 3.841 (df=1), 5.991 (df=2), and 7.815 (df=3) for ascending lengths of consonant clusters) indicating a preferential segmentation response. Those clusters/singletons which did not yield significant preferential responses were only found in the first condition, consisting of /ks/ and all single intervocalic consonants excepting /t/, /d/, and /l/.

For all clusters and singletons, bar one, there was agreement between the preferential segmentation of all stimuli within each particular feature category (FN, PF, etc.). There was also agreement between the preferential segmentations of these feature categories for each of the experimental conditions, first and second syllable repetition. The exception to both these cases was the OBLI (in this study the definition of OBLI clusters given by Dell in section 2.2.3 will be used unless otherwise indicated) cluster /vr/, whose preferential segmentation was /v.r/ for the first condition but /.vr/ for the second. This cluster was also at odds with the other found in the FL feature category, that of /zl/, whose preferential segmentation was /z.l/ in both conditions. However, it should be noted that /zl/ is not considered an OBLI cluster by Dell or Laporte.

Leaving this cluster aside for the moment, and examining the preferential segmentation of all other consonant clusters and singletons a clear pattern of segmentation for the stimuli was found. In single intervocalic stimuli, segmentation occurs before the consonant /V.CV/, for double consonant stimuli segmentation results in a single consonant onset, excepting OBLI clusters (/gr/, /br/, and /gl/ from the test data), which are segmented with a double consonant onset /V.CCV/. Triple consonant cluster stimuli are segmented with a double consonant onset for NPL and LPL feature categories, and a single consonant onset for that of PFP (/ksp/ and /kst/ clusters).

3.3.4.3 Syllabification Consistency

Other analyses, as important as the subjects' preferential segmentation response, are those concerning syllabification consistency. As can be seen in Figure 5, which covers all singular intervocalic consonant stimuli, the percentage of subject responses for the preferred boundary placement, that is the segmentation given by the majority of subjects, are higher for condition 2 than condition 1. For stimuli of greater complexity, such as the double intervocalic consonant cluster stimuli, shown in Figure 6, there are still differences in syllabification consistency between experimental conditions, but are restricted to a subset of feature categories. In more complex stimuli the syllabification consistency differences found between types of consonant cluster can be considerable. For example, taking results from the second experimental condition, some feature categories like NF have segmentation consistencies of 98.5%. Others, such as PF, have only 63.6%.

3.3.5 Discussion of Experiment 1

3.3.5.1 Syllabification Consistency in Singular Intervocalic Consonant Stimuli

As noted in section 3.3.4.3, there is a considerable difference in syllabification consistency of singular intervocalic consonants between experimental conditions 1 and 2, this finding is of particular interest as there is no theoretical controversy surrounding the syllable boundary placement in these cases, with segmentation following the Obligatory Onset Principle. If subjects follow this principle then responses should show a segmentation of /V.CV/ for both experimental conditions. However, an analysis of results show a considerable percentage of ambisyllabic response for these stimuli, that is, when a subject uses different segmentation for the same stimuli in Condition 1 and 2 such that a segment is common to the first and second syllables. The distribution of ambisyllabic responses across all singular intervocalic consonant categories can be seen in Figure 7. These results appear to be consistent with Kahn's (1976) theory of ambisyllabicity, which suggests that medial consonantal segments can be part of two syllables.

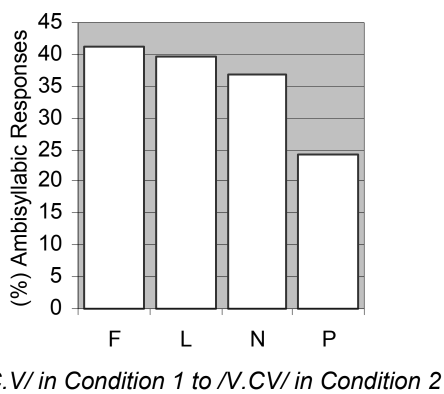


Figure 7: *Percentage of ambisyllabic response for /VCV/ stimuli in Experiment 1*

However, whilst Kahn's theory can explain the existence of ambisyllabic responses it does not explain the large disparity found between syllabification consistency in experimental conditions, with over 99% consistency in condition 2 and only 66% in condition 1. With almost all responses in condition 2 resulting in a single consonant onset the ambisyllabic responses are due to a change in segmentation from /VC.V/ in condition 1 to /V.CV/ in condition 2, whilst only 3.5% ambisyllabic responses had the reverse behaviour (/V.CV/ in condition 1 to /VC.V/ in condition 2).

Similar differences in syllabification consistency between the repetition of the first and second parts of stimuli were also found in an unspeeded study of syllable boundary placement in single intervocalic consonant stimuli (Content et al, in press). In the study of Content et al (in press) this effect is cited as evidence against the boundary view of syllable segmentation. Instead, an 'onset hypothesis' is suggested in which distinct operations are used to detect syllable onsets and offsets, with the former providing reference points for segmentation and lexical access. For first syllable repetition it is suggested that the syllable boundary is set by the detection of the end of that syllable, its offset. However, in second syllable repetition the boundary is set by the detection syllable's onset.

If this hypothesis is correct, and there are different mechanisms for the detection of syllable onsets and offsets, this should be reflected in the distribution of syllabification consistency across different consonant clusters/singletons. If no positive correlation is found between the consistency of syllable onset and offset placements, this might signify that consonant clusters/singletons are treated differently dependent upon which are used for segmentation, the onset or offset. However, if a positive correlation is found then it is less likely that distinct operations were used for the detection of syllable onsets and offsets, as different consonant clusters/singletons are treated in the same manner by both operations. To test this supposition the percentage of subject responses for the preferential segmentation of each of the singular intervocalic consonant stimuli were compared. The results showed that there was no significant correlation in syllabification consistency ($r = -.366$, $t(8) = -1.112$, $p < 0.05$) between onset and offset detection. This result, along with the disparity in syllabification consistency found between experimental conditions, suggest that results of Experiment 1, at least for singular intervocalic consonant stimuli, uphold the tenets of the onset hypothesis.

In addition to those found between experimental conditions, other syllabification consistency disparities were found between different stimuli. It is suggested (Treiman & Danis, 1988) that, for single intervocalic stimuli, sonorants are more closely linked to the preceding vowel than obstruents. This is taken as evidence for the use of sonority, as this behaviour fits with the preferred contour of a steep sonority rise up to the nucleus of a syllable, and slow decline after. Examining subject responses from the first experimental condition there are similar differences in /VC.V/ responses for plosive and liquid categories (40.8% of responses for Liquid clusters, 24.4%

for Plosive) to those found by Treiman and Danis (1988) using syllable reversal and orthographic tasks. For responses of the second experimental condition this effect was not found with 0.8% /VC.V/ response for both categories. However, if segmentation is performed by syllable onset detection, as suggested by the onset hypothesis, then the evidence for the use of sonority is moot as it is only found in syllable offset detection, which would be of little importance to syllabic segmentation. Unfortunately, the experimental tasks presented in the Treiman and Danis (1988) could not show the disparity between onset and offset detection.

3.3.5.2 Differences in Syllabification Consistency between Consonant Clusters

A cursory examination of segmentation responses for double intervocalic consonant clusters, shown in Figure 6, shows that the distribution of syllabification consistency is far more complex than that found with singular intervocalic consonants. Whilst some consonant clusters have relatively high consistency for both experimental conditions (like LN clusters), for others there is a disparity in syllabification consistency between experimental conditions, with some (such as FP clusters) having higher consistency in Condition 1 than 2, and others the opposite (such as PL clusters).

Two possible hypotheses are proposed to explain the differences in consistency amongst consonant clusters. The first of these is based upon the relative frequency of occurrence of the various consonant clusters, and the second, upon the availability of multiple legal onsets. The first suggests that subjects are able to segment those consonant clusters that are relatively common in the language without difficulty, whilst encountering problems those with which they have less familiarity. In this case consonant clusters with a high frequency of occurrence will have higher segmentation consistency than those of lower frequency. The second hypothesis is based upon the Law of Initials of the Legality Principle. This states that if, for a particular consonant cluster, there are multiple legal onsets available then, because of the choices offered in syllable boundary placement, that consistency would be lower than if only one legal onset were available.

To test the first hypothesis a measure of the frequency of occurrence has to be found for each of the double consonant clusters. This was produced by searching BRULEX (Content, Mousty & Radeau, 1990), a phonetic lexicon containing word frequency information, with the frequency of occurrence of each consonant cluster calculated as the sum of all word frequencies containing that cluster. The natural log of the frequency of each of the 40 clusters was then compared with its syllabification consistency using tests of correlation (with 40 data points, one for each cluster) for each experimental condition. The syllabification consistency of each cluster was calculated by selecting the preferred syllabification decision (that with the highest number of subject responses) and dividing the number of responses given for this particular syllabification decision by the total number of responses. For example, for the cluster /ft/ there were 36 subject responses for the syllabification /f.t/ and 8 for /.ft/, in this case the preferential response is /f.t/ with a syllabification consistency of 0.81 (36/(36+8)). If the hypothesis is correct there should be a significant positive correlation between these factors, indicating the relationship between consonant cluster frequency and syllabification consistency. The results showed that there were no significant ($p>0.05$) positive correlations for the first ($r = -0.456$) or second ($r = -0.26$) experimental conditions. In these cases there appears to be no evidence to support the first hypothesis.

To verify the second hypothesis a search of BDLEX gave the number of possible legal onsets, excluding those not in agreement with the Obligatory Onset Principle (those with a zero consonant onset), for each of the double intervocalic consonant clusters. (e.g. the cluster /ft/ has two legal onsets, /ft/ and /t/, whilst the cluster /mv/ has only one, /v/). A full list of all these clusters can be found in Appendix B, along with the number of legal onsets associated with each (1 or 2). As in the test of the previous hypothesis, a test of correlation was made between the syllabification consistency of each of the 40 double intervocalic consonant clusters, and, in this case, the number of the number of possible onsets for each cluster. As before, one correlation (each with 40 data points) was used for each of the experimental conditions. If the hypothesis is correct there should be a significant negative correlation between these factors, indicating the inverse relationship between availability of multiple legal onsets and segmentation consistency. For responses from the first experimental condition there was no significant negative correlation ($r = -0.11$, $t(39) = -.692$, $p > 0.05$) between the two factors. However, for responses from the second experimental condition there was a significant negative correlation ($r= -0.656$, $t(39) = -5.423$; $p<0.05$). These findings are also reflected in average syllabification consistency, calculated using subject responses from experimental condition 2, as double intervocalic consonant clusters with two legal onsets had lower average consistency (80.6%) than those with only one possible onset (94.7%).

This result reveals two possible aspects of syllabification. Firstly the importance of phonotactic regularities, embodied by the Law of Initials of the Legality Principle, on the subject's syllabification responses using syllable onset detection, and secondly, the differences in syllabification between experimental conditions.

3.3.5.3 Ambisyllabicity between Experimental Conditions

The findings of the previous section revealed a link between syllabification consistency for subject responses in experimental condition 2, and the number of legal onsets available for a particular consonant cluster. These results suggest that the influence of the law of initials on syllabic segmentation is complex, with the factors of both experimental condition and the type of consonant cluster involved in its relationship with syllabification consistency. Therefore, in order to simplify the examination of these relationships, subject responses have to be broken down into categories, those that remain constant across experimental condition, and those that are variable.

In Experiment 1 each stimulus was presented to the subject twice, once for each of the experimental conditions, and, for a double intervocalic consonant cluster stimulus, there are three possible segmentation responses, /.CC/, /C.C/, and /CC./. Therefore, if all possible combinations of segmentation responses are considered across *both* experimental conditions there are nine categories of response (/.CC/ in condition 1 and /.CC/ in condition 2, /.CC/ in condition 1 and /C.C/ in condition 2, etc.) for each stimulus. Unfortunately, a detailed examination of all nine categories would be rather complex. Therefore, to simplify matters, the five categories involving the /CC./ segmentation will not be used in further analyses. As only 3.5% of all segmentation responses fall into these categories they are unlikely to have significant impact on syllabification consistency. Furthermore, a large proportion (over 1/3) of /CC./ responses are for the cluster /ks/, which has special, orthographic considerations, a subject which will be examined later.

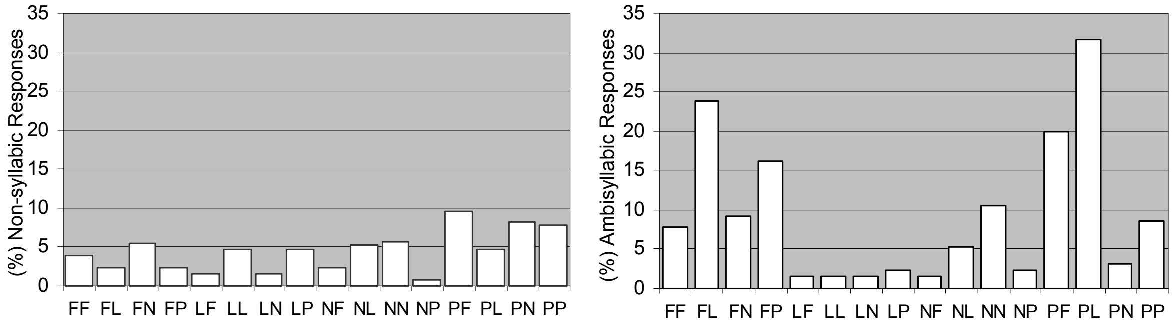
The four remaining categories can then be split into two groups, the first describes the two categories where the subject responds with the same segmentation for each condition (*double consonant onset category*, /.CC/ in condition 1 and /.CC/ in condition 2, and *single consonant onset category*, /C.C/ in condition 1 and /C.C/ in condition 2). In the remaining categories there is an *ambisyllabic* response, that is, a different segmentation response was given in the two experimental conditions, which, taking the segmentation for both conditions together, results in an ambisyllabic consonant (e.g. /p.s/ in Condition 1 and /.ps/ in Condition 2, with [p] as the ambisyllabic consonant). For the sake of convenience, the ambisyllabic categories will be referred to by relative length of the syllable onset between experimental conditions. As onset detection has been linked to experimental condition 2, the responses will be referred to as *reduced onset ambisyllabic responses* (/.CC/ in condition 1 and /C.C/ in condition 2), and *increased onset ambisyllabic responses* (/C.C/ in condition 1 and /.CC/ in condition 2).

To estimate syllabification consistency, it is necessary to have an understanding of the factors that affect the distribution of subject responses in each of the four categories. For two of the categories described previously, that is, the *double consonant onset category,* and *single consonant onset category,* the segmentation remains constant across the two experimental conditions. That is, the subject gave the same segmentation decision for a particular stimulus when repeating the first and second parts of the stimulus (conditions 1 and 2). Therefore, if there is a /.CC/ segmentation in the condition 1, the subject also segmented that cluster as /.CC/ in condition 2 (*double consonant onset category)*. These categories of response are unaffected by the issues related to the differences found in syllabification consistency between conditions, that is, syllable onset and offset detection.

These categories are also those that contain the largest percentage of subject responses (82% of all responses) and as such, are usually the categories that will decide the preferential segmentation response of each cluster (the segmentation response with the majority of subject decisions).

The remaining categories, containing ambisyllabic responses, represent the differences in syllabification due to experimental condition, highlighting the differences between syllable onset and offset detection. Whilst subject responses in the first two categories were influenced by the same factor, with one category showing the inverse distribution of responses to the other, it is unlikely that the same can be said of the ambisyllabic categories.

A cursory examination of the two categories of ambisyllabic response, shown in Figure 8, reveals considerable differences in the distribution of responses across different consonant clusters. Whilst reduced onset ambisyllabic responses are relatively evenly distributed amongst consonant clusters (average percentage all of responses is 4.5%, S.D. is 3.34%), the distribution of increased onset ambisyllabic responses is highly dependent upon the type of consonant cluster (average percentage of all responses is 9.4%, S.D. is 10.7%).



/V.CCV/ in Condition 1 to /VC.CV/ in Condition 2 *VC.CV/ in Condition 1 to /V.CCV/ in Condition 2*

Figure 8 *Percentage of ambisyllabic responses for /VCCV/ stimuli in Experiment 1*

In the previous section it was found that consonant clusters with two legal onsets had lower syllabification consistency than those with one, but only for responses from the second experimental condition. This suggests that subject syllabification responses were constrained by the law of initials, but only for syllable onset detection. Therefore subjects tend to respond with double consonant onset responses only for consonant clusters where the response is legal, whilst single onset responses are legal for all the double consonant stimuli. This behaviour is reflected in the distribution of ambisyllabic responses. For reduced onset ambisyllabic responses the constraints of the law of initials do not come into play as the /C.C/ response in Condition 2 is legal for all clusters (there are no constraints on the responses of Condition 1 for /V.CCV/ and /VC.CV/ segmentations). This results in a relatively even distribution of these responses over all consonant clusters. However, for increased onset ambisyllabic responses the law of initials constrains the V.CCV response in Condition 2 to those clusters where a double consonant onset is legal. Therefore, those consonant clusters with two legal onsets (such as PL clusters) should have a higher incidence of increased onset ambisyllabic response compared to those with only one legal onset (such as NF clusters).

A more rigorous investigation of the possible link between the number of legal onsets and increased or decreased onset ambisyllabicity was conducted by making a test of correlation between these factors. Two separate tests of correlation were conducted, one between the number of legal onsets available for each of the 40 double consonant clusters and the number of responses resulting in increased onset ambisyllabic behaviour, and the other for responses resulting in decreased onset ambisyllabic behaviour. For each of the 40 double consonant cluster stimuli the number of legal onset available for that cluster was compared with either increased of decreased onset ambisyllabicity. Increased onset ambisyllabicity was calculated by comparing syllabification responses from experimental conditions 1 and 2, and counting the number of syllabification responses resulting in /C.C/ segmentation in condition 1 and /.CC/ in condition 2. Similarly, the number of decreased onset ambisyllabic responses was calculated from responses that resulted in /.CC/ segmentation in condition 1 and /C.C/ in condition 2.

It was found that these analyses showed no significant correlation between decreased onset ambisyllabicity and the availability of multiple onsets ($r = .185$, $t(39)=1.177$), but a significant positive correlation between multiple legal onsets and increased onset ambisyllabic responses ($r = .799$, $t(39)=8.308$, $p<0.05$).

These new findings clarify the true extent of the influence of the Law of Initials on syllabification consistency, and require an amendment to the conclusions of the original analysis of the link between the two. The first analyses showed a significant negative correlation between the number of legal onsets available for a double consonant cluster and syllabification consistency for syllable onset detection, but not offset detection. This predicts lower syllabification consistency for responses from experimental condition 2 in clusters with two legal onsets than those with only one. However, the new findings suggest that the influence of the Law of Initials is limited to increased onset ambisyllabic responses, with a greater number of these responses for clusters with multiple legal onsets. This means that the relationship between the number of legal onsets for a consonant cluster and syllabification consistency is also dependent upon the ratio of subject responses in the single and double consonant onset categories.

The possible relationship between the subject responses in each of the four categories and syllabification consistency is outlined in Figure 9. To calculate syllabification consistency, the percentage of subject responses for the preferential segmentation, it is necessary to find the number of responses for each of the two segmentation responses (/.CC/ and /C.C/) for each experimental condition. These can be found by adding the number of responses in each of the categories with the same segmentation for each condition. For example, the number of /.CC/ responses in condition 1 is the sum of the responses in the double consonant onset category and the reduced onset ambisyllabic category (each have a /.CC/ segmentation in Condition 1). The syllabification consistency for each condition is the ratio of the number of responses for the segmentation with the most responses, over the number of responses found in both.

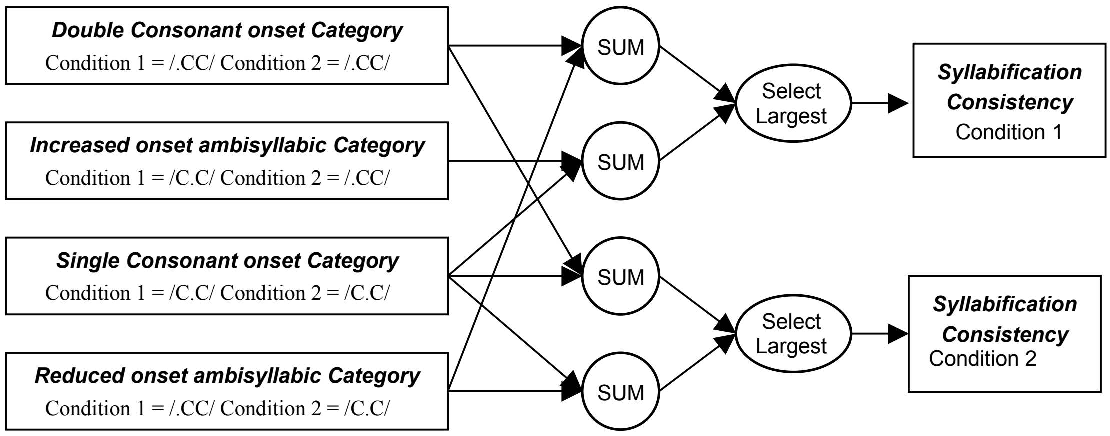


Figure 9 *Calculation of syllabification consistency from cross condition segmentation categories*

Therefore, for the majority of double intervocalic consonant clusters, those where the majority of subject responses are in the single consonant onset category, multiple legal onsets lead to lower syllabification consistency in condition 2. However, for the other, OBLI clusters, where the majority of subject responses are in the double consonant onset category there is higher syllabification consistency in condition 2.

These findings are also reflected in the average syllabification consistencies calculated across all double intervocalic consonant clusters for the three groups of consonant cluster. Clusters with a single consonant onset have high average syllabification consistency for both conditions, at 92% for Condition 1 and 94% for Condition 2. For consonant clusters with two legal onsets, those with a /VC.CV/ preferential response have 86% average syllabification consistency for Condition 1, and 77% for Condition 2, for those with a /V.CCV/ preferential response it is 65% for Condition 1 and 88% for Condition 2.

Therefore, it appears that it is possible to predict differences in syllabification consistency between experimental conditions with some accuracy, based on whether the cluster is an OBLI, and how many legal onsets it has. These findings also help explain the preferential segmentation of the /vr/ OBLI cluster, segmented as /v.r/ in experimental condition 1, and /.vr/ in condition 2. For this cluster the percentage responses in the double consonant onset category (37%) is not much higher than that found in the single consonant onset category (19%). Therefore, with 35% of responses in the increased onset ambisyllabic category, the preferential segmentation reflects the segmentation behaviour of this category, with /VC.CV/ segmentation in condition 1 and /V.CCV/ in condition 2.

3.3.5.4 Conclusions

Thus far a number of factors have been suggested as influences upon syllabic segmentation. An examination of the preferential segmentation of intervocalic consonant clusters and singletons under analysis suggest that, in the majority of cases, segmentation results in a single consonant onset. The only exceptions involve OBLI clusters, where the syllable boundary is placed before the cluster. However, other results suggest that the preferential responses of experimental subjects do not show all of the influences on syllabic segmentation. An examination of syllabification consistency, as well as ambisyllabic responses, shows differences in segmentation between experimental conditions, and amongst different consonant clusters/singletons.

It has been suggested by Content et al (in press) in the ‘onset hypothesis’ that different processes are used for the detection of syllable onsets and offsets, with syllable onset detection used for segmentation and lexical access. Evidence for this hypothesis comes from syllable repetition tasks with singular intervocalic consonant stimuli, which found higher syllabification consistency for second syllable repetition than first syllable repetition. Analyses of subject responses for singular intervocalic stimuli in Experiment 1 also appear to support this hypothesis, with a similar disparity found in syllabification consistency between the two experimental conditions. Additional evidence comes from an examination of the similarities between syllabification consistency in conditions 1 and 2 across all singular intervocalic consonant stimuli. This found that there was no significant correlation between the two, which would have been expected if there were separate processes for syllable onset and offset detection.

However, an examination of the distribution of segmentation responses for double intervocalic consonant cluster stimuli did not show the same clear distinction between syllabification consistency for syllable onset and offset detection. Analyses suggest that syllabification consistency is dependent upon the number of increased onset ambisyllabic responses (/VC.CV/ segmentation in Condition 1 and /V.CCV/ in Condition 2). For clusters with one legal onset these responses are balanced by those of reduced onset ambisyllabic responses (/V.CCV/ segmentation in Condition 1 and /VC.CV/ in Condition 2), resulting in high syllabification consistency in both experimental conditions. For clusters with multiple legal onsets there are a higher proportion of increased onset ambisyllabic responses than reduced onset ambisyllabic responses. So, for clusters with a single consonant onset this results in a syllabification consistency disparity between experimental conditions, higher in Condition 1 than 2. For clusters with a double consonant onset the disparity is reversed. Therefore, taking the results of the analyses of syllabification consistency for both single and double intervocalic consonant clusters, syllabification consistency is only higher in condition 2 than 1 when preferential segmentation results in a syllable boundary adjacent to a consonant cluster/singletons' preceding vowel.

In the light of these findings, with syllabification consistency higher for syllable onset, rather than offset, detection for only a small minority of stimuli, it seems unlikely that syllable onset detection has special status in syllable segmentation as suggested by the onset hypothesis. However, with the relationship between the segmentation behaviour and the Law of Initials limited to responses from experimental condition 2, it still appears that there is evidence for the separation of the processes of syllable onset and offset detection. In order to explain the differences in syllabification consistency between these processes a new hypothesis is tentatively suggested, the 'ambisyllabic onset hypothesis', based upon Kahn's (1976) theory of ambisyllabicity, the onset hypothesis (Content et al, in press), and the Law of Initials of the Legality Principle. This hypothesis suggests that if there is a legal syllable onset before the first consonant in an intervocalic consonant cluster/singleton then there will be a high degree of increased onset ambisyllabicity for that consonant. This predicts higher syllabification consistency for syllable onset detection than offset detection when the syllable boundary of the preferential segmentation lies before the first consonant, with the disparity reversed if the preferential response lies behind the consonant.

Whilst the 'ambisyllabic onset hypothesis' can predict the occurrence of syllabification inconsistency with a degree of accuracy, there is one influence, touched upon briefly in Section 3.3.5.2, which has not been taken into account, that of orthographic bias. Because the task used in Experiment 1 was a metalinguistic, not 'on-line', task then it is possible that metalinguistic knowledge, such as that used in orthographic tasks, may influence subject responses. In an attempt to avoid such contamination non-word stimuli were used in the experiment. However, it appears that, at least for stimuli containing the /ks/ cluster, subject responses are affected by orthographic constraints. The /ks/ cluster is unusual in that it is most often represented by a single grapheme, 'x', as in the word '*klaxon*', while its bi-grapheme forms are less common ('cc' or 'cs'). It is thought that this representation can cause ambiguity as to whether the cluster contains one or two segments. Evidence of the confusion surrounding the segmentation of this cluster can be seen in the distribution of subject responses in experimental condition 1. In section 3.3.4.2 it was noted that, for subject responses from condition 1, there was no significant preferential segmentation response for /ks/ (the only cluster without a significant preferential segmentation), with the distribution of responses being 21% for /.ks/, 31% for /k.s/, and 48% for /ks./. The distribution of subject responses is unlike any other consonant cluster tested in Experiment 1, in that there is an extremely high proportion of zero onset responses, showing that the majority of subjects believe that the /ks/ cluster is indivisible. This result is unlike the distribution of segmentation responses from both of the other clusters in the PF category, which have the significant preferential segmentation response /VC.CV/ for Condition 1, with an average syllabification consistency of 78%. As both of the other clusters in PF can be represented by double graphemes (although /gz/ is also represented by the grapheme 'x', this is less frequent, it is more often by 'cs' or 'gz') the evidence appears to point towards orthographic bias as the influence behind the distribution of subject responses for the /ks/ cluster in experimental condition 1. However, an examination of the distribution subject responses from condition 2 shows no evidence for orthographic bias in the syllabification of the /ks/ cluster. In condition 2 there were no zero onset responses for /ks/, with a significant preferential segmentation of /VC.CV/ for all clusters in the PF category. There was also very little difference in syllabification consistency between /ks/ and the other PF clusters, with 64% consistency for /ks/ and an average of 63% for the other two PF clusters.

These findings suggest that subject responses, at least those of experimental condition 1, are subject to influence due to the use of orthographic knowledge. In the case of the cluster /ks/ this influence is clear. However, this finding also highlights the possibility that orthographic rules may play broader role in the segmentation responses given by the subjects in Experiment 1. If the preferential segmentation responses given across the whole range of stimuli are being influenced, or even decided, by 'learnt' orthographic rules, rather than unconscious segmentation, then the results of Experiment 1 would be invalid. In this case the results would merely be a distorted reflection of these rules and not an insight into pre-lexical language processing, as was intended. Perhaps the simplest method of validating Experiment 1 would be to compare its results with those from an experiment that was free from orthographic contamination. There are two methods by which this may be achieved, firstly, the use of an 'on-line' task, which should not allow the use of lexical processes, and secondly, the use of illiterate subjects, who have no knowledge of orthographic rules. Unfortunately, on-line tasks, using the 'syllable effect' (Mehler et al, 1981) to test the location of the syllable boundary are highly problematic, as this effect can be highly elusive (Meunier et al, 1997). The simpler solution would be to use an experiment similar to that of Experiment 1, but using illiterate subjects.

3.4 The Role of Orthographic Knowledge in Syllabic Segmentation[9]

The main aim of the experiments presented in this section is the evaluation of the role of orthographic knowledge in the responses of experimental subjects in metalinguistic syllable segmentation tasks. To achieve this aim the syllabification behaviour of two subject groups will be compared, those of pre-literate children, and literate adults. Children of around 5 years are thought to have acquired the knowledge required for syllabic segmentation (as discussed in Section 1.3), but have yet to learn the orthographic rules used for reading and writing. Therefore, a comparison between the syllabification behaviour of children and adults should provide some indications of the role of orthographic knowledge in this task.

It is acknowledged that any findings resulting from a direction comparison of experimental data from pre-literate children and adults must be taken with caution as there are many factors, as well as the acquisition of orthographic knowledge, that differ between the two experimental populations. At four or five years of age the development of cognitive processes is incomplete, and it is expected that continuing development will improve the performance of experimental subjects in experimental tasks that test metaphonological abilities. For example, short-term memory span continues to increase until early adolescence (Pascual-Leone, 1970), a factor that could have limited the efficacy of the syllable repetition task in subjects from early childhood.

However, previous studies have shown that children of around five years old are capable of performing experiments designed to elicit syllabification decisions, at least in a “tapping” task, where subjects repeat the stimuli slowly, inserting a pause between each syllable (see Liberman et al., 1974, for a use of this task with children). This ability is possibly due to the nature of the task, which does not require a high mental load, and resembles a number of common word games. In addition, in such a task the only metaphonological ability that is required does not touch upon phonemic awareness, but only syllabic awareness. Syllabic awareness, as with rhythmic awareness, is supposed to be available from a very early age, contrary to that of phonemic awareness (Jiang & Peng, 1999; Goswami & Bryant, 1990; Alegria & Morais, 1979). Finally, touching upon the

[9] The experiments presented in Section 3.4 have been submitted for publication in a collaborative investigation of the developmental aspects of syllabic segmentation (Floccia & Goslin, submitted), made with Dr. Caroline Floccia at the University of Franche-Comté, France. Whilst the developmental aspects of syllabic segmentation are not of direct relevance to this thesis, a subset of the experiments carried out by the author in this collaborative investigation are of relevance to this thesis. These experiments examine the roles of orthographic and morphological knowledge in metalinguistic tasks by comparing the syllabification responses of pre- and post-literate subjects.

differences in language exposure, whilst children have far shorter exposure to language than adults, as was mentioned in section 1.3, it is often acknowledged that the basic knowledge of syllabification rules is acquired at 5 years (Fallows, 1981; Gillis & De Schutter, 1996). These arguments suggest that four to five years old children should exhibit a similar level of competence in syllabification tasks as adults, but with the important exception that adults have access to learned orthographic rules. This knowledge may trigger their metaphonological knowledge and bias their syllabification representations in psycholinguistics tasks.

Responses will be gathered from both subject groups using the metalinguistic 'tapping' task, where subjects repeat the stimuli slowly, inserting a pause between each syllable. This task replaces the repetition task used in Experiment 1 because it was thought, as mentioned earlier, that juvenile subjects would have greater familiarity with this task, since it has similarities with a number of common word games. This should increase subject attention, and so reduce experimental error due to uncertainty or lack of interest. An additional advantage is that, by introducing a new experimental paradigm, segmentation results from different tasks can be compared to ensure there is no contamination due to task specific influences.

In Experiment 2 lexical stimuli were selected to encompass a wide range of double intervocalic consonant clusters. It must be noted that this experiment was not designed to test specific orthographic effects (such as that noted with the cluster /ks/ in the previous experiment), but to show the extent of the influence of orthographic knowledge over a similar range of stimuli to that used in Experiment 1. In Experiment 3, the lexical stimuli used in the previous experiment were replaced with non-words. This change was made to investigate the possibility of morphological influences upon the syllable segmentation of words, and also to verify that the preferential segmentation behaviour found in Experiments 2 and 3 was valid over a wider range of stimuli.

3.4.1 Experiment 2 – Syllable Tapping Task using Lexical Stimuli with Adults and Pre-literate Children

3.4.1.1 Stimuli

Test items

As in Experiment 1, consonant phonemes are divided into four classes, Fricatives, Plosives, Nasals, and Liquids, resulting in sixteen possible double intervocalic consonant categories. Unfortunately, in using lexical stimuli only 11 of these 16 categories could be used. The five remaining clusters were avoided because they were only found in rare (FF as in *asphalte*) words that children would not know, or because they are spelled with an 'e' (NN as in *promener*) and would obviously lead to orthographic ambiguity. Where possible the same double intervocalic consonant clusters used in Experiment 1 were also included in Experiment 2. However, for the reasons outlined above, it was not always possible to find suitable lexical stimuli for all clusters, therefore only a subset of clusters used in Experiment 2 will be available for direct comparison with Experiment 1.

For six out of the eleven cluster types selected (PL, LP, LN, LF, FP, FL), 16 frequent bi-syllabic French words were selected using their phonetic transcription and frequency information found in the BRULEX French lexicon. The five remaining cluster types (PF, LL, PP, PN, FN) were not sufficiently frequent to allow the selection of 16 words per consonant category, thus a smaller number of stimuli were used. In total, 48 different double intervocalic consonant clusters were used in the set of stimuli. When selecting stimuli an attempt was made to include as many different tokens as possible for each cluster type. A list of all stimulus words can be seen in Appendix C.

Stimuli were produced by 37 different monolingual female speakers from the Franche-Compté region of France. The speakers were instructed to produce the words from a randomised list of words with unchanging intonation contours, and with a pause of three seconds inserted between the production of each stimulus. Stimuli were recorded and presented to subjects over headphones or loudspeakers. The implication of this multiple speaker design is that it should help to generalise any experimental effects due to speaker variability. However, it must be noted that, to avoid confusion in the subjects, that each subject was tested with only one speaker's voice.

Distractor and Training items

A total of 36 distractor items were used in the tapping task, each contained between one and three syllables. Distractor items were selected using a number of constraints to ensure there was similarity between distractor and test items. These state that there should be an equal number of distractors starting with vowels and consonants, that at least 30% be infinitive verbs, and that they could not contain any of the consonant clusters used in the test items.

A total of 10 training items were used, these were selected using the same criteria as the distractor items.

3.4.1.2 <u>Procedure</u>

In Experiment 2 syllable segmentation responses were collected by means of a tapping task, in this task subjects repeat each word slowly and insert a pause between each syllable. This task is called tapping because it was originally designed by tapping on the table while repeating every syllable (Liberman et al, 1974). In the present study, subjects were also invited to add rhythm to their speech by clapping their hands.

Subjects were tested individually in a quiet room. Children were asked to participate in a short game involving an alien puppet that spoke "like a robot", segmenting all the words into syllables (*Bon-jour-je-suis-le-ro-bot*). It wanted to learn new words, but understood only if it was spoken to like a robot as well. Children were invited to help it in learning new words by repeating the recorded ones and adding a pause between each 'part' of the word. When testing adults the procedure divested in the use of the puppet, adult subjects were simply invited to repeat the word slowly by introducing a pause at each segment. A few examples with feedback were given before each task. The word syllable was not used, and it appeared that with a few simple examples, no further explanations were needed for most subjects.

In order to keep the experimental duration to a minimum, suitable for use in testing juvenile subjects, each subject was randomly assigned a block of stimuli, chosen from multiple cluster types and not exceeding one eighth of all experimental stimuli (16 stimuli).

3.4.1.3 Participants

Two subject groups were tested, pre-literate children and literate adults. In the first group there were 137 subjects (67 boys and 70 girls) with an average age of 4;4[10](ranging from 3;4 to 5;9). All juvenile subjects were at the same level of schooling, approximately one and a half years before the tuition of reading and writing. It was not possible to guarantee that all subjects had no orthographic competences, as individual parents could have initiated separate tuition. However, as the experiment was performed in schools it was possible to interview the children’s teachers in an attempt to ascertain whether individuals had noticeable skills in these areas, those that did were not selected for the experiment. The data of 58 additional children were discarded for the following reasons: lack of subject response (18), lack of interest or concentration (11), experimental error (1), misunderstanding of the task and/or error rate on test items higher than 50% (28).

In the second subject group there were 147 subjects (66 men and 81 women) with an average age of 25;2 years (ranging from 17 to 53). Two subjects were rejected, one because of high error rates (the subject systematically gave a CV response to all stimuli) and the other because of hesitation.

All subjects were native French monolingual speakers, and had no recorded auditory problems. As with the speakers that produced the experimental stimuli, both child and adult participants were natives of Franche-Comté. This is a region of France that borders with Suisse Romande, with both sharing a similar regional dialect (with only minor differences, such as with the

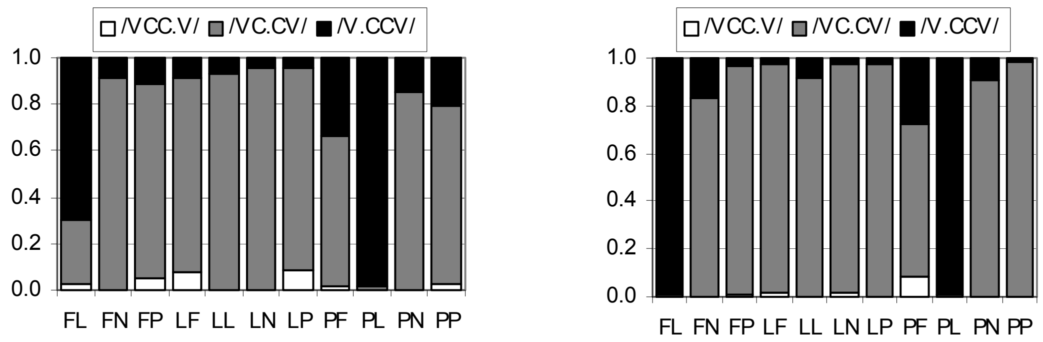
[10] Ages are denoted using years;months (e.g. 4;4 is 4 years and 4 months).

location of word accent). The phonemic inventory of both regions is very similar, especially concerning the close/open vowels distinction. For instance, the word “jeune” (young) in “neutral” Parisian French will be pronounced as /ʒœn/ and /ʒøn/ in Geneva or in Franche-Comté.

3.4.1.4 Results

Error rates

Of the 1937 responses collected with pre-literate subjects 7.67% were in error. For adult subjects 1.12% of the 2326 responses were erroneous. Errors included subject uncertainty, experimenter uncertainty, tapping errors (where the subject failed to tap when presented with a stimulus, or tapped more than once), and repetition errors. There was no significant effect of cluster type on error rates for either adult or pre-literate subjects.



Pre-literate Subjects ***Literate Subjects***

Figure 10 *Distribution of subject syllabification responses in Experiment 2 for pre-literate children and literate adults for all robust feature classes of stimuli*

Preferential Segmentation

Segmentation responses were analysed separately for each of the 48 cluster types. For each cluster type a χ^2 was calculated on the three possible responses, double consonant syllable onset (V.CCV), single consonant syllable onset (VC.CV), or zero onset (VCC.V), to determine whether there was a significant preferential response. In each case, the χ^2 was higher than the critical χ^2 at p=.05 (5.991 (df=2)), indicating a preferential segmentation response. The pattern of results shows that all clusters were segmented with a single consonant onset, apart from those in the FL and PL categories, all OBLI clusters, which were segmented with double consonant onsets.

3.4.1.5 Discussion of Experiment 2

In the discussion of Experiment 1 it was suggested that the use of a metalinguistic syllabification task with literate subjects led to contamination of results due to the influence of orthographic knowledge on subject responses. One of the primary aims of Experiment 2 was to examine the extent of this influence, whether it was simply limited to specific clusters, such as /ks/, or had a wider effect, responsible for the general syllabification strategies used by the experimental subjects. To achieve this aim syllabification responses, collected by means of a tapping task, were taken from two subject groups, one consisting of pre-literate children, who should have no knowledge of orthographic rules, and adult, literate subjects.

A comparison of the preferential segmentation and syllabification consistency for the two groups suggest that the influence of orthographic knowledge upon metalinguistic syllabification tasks is relatively minor, as similar segmentation behaviour was used in both subject groups. For both subject groups preferential segmentation resulted in a single consonant onset, excepting OBLI clusters, which were segmented with double consonant onsets. Syllabification consistency results, shown in Figure 10, also reveal considerable similarity between the segmentation behaviour of subjects in the two groups. Whilst the pre-literate subjects show lower average syllabification consistency than adults, with 85% consistency for the former and 93% for the latter, syllabification consistency differences amongst consonant clusters are similar, with a significant correlation found between the distribution of syllabification consistency for the two groups ($r=0.573$, $t(47)= 4.789$, $p<0.05$). It is suggested that the lower overall syllabification consistency in the pre-literate child group is more likely to be due to general uncertainty concerning the experimental task, rather than an aspect of syllabic segmentation at this age. This is reflected in the relatively high error rates amongst children, with 7.4% of responses discarded due to errors caused by the subject (subject uncertainty, tapping errors, and repetition errors), compared to only 1% for adult subjects.

The only cluster where subject responses were clearly affected by the use of orthographic knowledge is that of /ks/. In Experiment 1 it was suggested that, because this cluster is normally represented by a single grapheme, ‘x’, that subjects were confused as to whether one or two segments represented this cluster. This resulted in a high proportion of zero onset responses (45% of responses) in experimental condition 1, indicating confusion amongst subjects and a reluctance to split this cluster between two syllables. In Experiment 2 the distribution of adult literate subjects show a similar behaviour to that seen in Experiment 1, with 11% of subject responses resulting in a zero onset segmentation. This pattern of responses is quite different to those of the remaining PF clusters, where there were only 1% of responses for zero consonant onset segmentation. In contrast, an examination of the responses from the pre-literate subject group reveals little difference between the percentage of zero consonant onset responses of /ks/ and the other PF clusters. Only 2% of segmentation responses for the /ks/ cluster resulted in a zero consonant onset, and none for the remaining PF clusters. A comparison of the frequency of zero onset responses between adult and child subjects in Experiment 2 was conducted using a χ^2 analysis. This analysis compared the number of subject responses for /.CC/ and /C.C/ with those of /CC./ for the two subject groups. This analysis found that there was a significant effect of subject group on the distribution of zero onset syllabification responses ($\chi^2 = 3.99$, df = 1, $p<0.05$).

These results appear to support the hypothesis put forward in the discussion of Experiment 1. Because the confusion surrounding the segmentation of /ks/ was only found in literate subjects, not in pre-literate children, it is likely that the unusual segmentation responses for this cluster are simply due to the interference of orthographic knowledge in literate subjects segmentation behaviour. However, as this finding is only supported by the observations from a single cluster, it must be treated with some caution.

Although the primary aim of Experiment 2 was the examination of the effects of orthography upon syllabification, the overlap of double intervocalic consonant clusters used in the stimuli of Experiments 1 and 2 also allows a superficial comparison of the syllable repetition and tapping tasks used in these experiments. Of the 41 double intervocalic consonant clusters used to generate the stimuli in Experiment 1, and the 48 used in Experiment 2, there were 19 clusters tested in both experiments. These clusters were used to compare both the preferential segmentation responses and syllabification consistency between experimental tasks.

As has been discussed, the preferential segmentation given by subjects in the two experiments are much the same, with the only differences those found between conditions 1 and 2 of Experiment 1. In Experiment 1 the preferential segmentation of the cluster /vr/ was /v.r/ in condition 1, but /.vr/ in condition 2. In Experiment 2 the preferential segmentation of this cluster is in agreement with condition 2, segmented with a double consonant onset, as was the case for all OBLI clusters. A comparison of syllabification consistency between experimental tasks also showed that there were greater similarities between the syllabification responses from the tapping task and condition 2 of the repetition task, than condition 1. A comparison of syllabification consistency for both pre-literate child and adult subject groups in Experiment 2 and condition 1 of Experiment 1 found no significant correlation between the two (for pre-literate child subject group $r=0.254$, $t(17)=1.082$, $p>0.05$, for adult subject group $r=0.381$, $t(17)=1.7$, $p>0.05$). However, the responses of condition 2 of Experiment 1 and syllabification consistency in both pre-literate child and adult subject groups in Experiment 2 both showed a significant correlation (for pre-literate child subject group $r=0.626$, $t(17)=3.307$, $p<0.05$, for adult subject group $r=0.541$, $t(17)=2.649$, $p<0.05$).

This finding is also supported by the predictions made by the 'ambisyllabic onset hypothesis'. For second syllable repetition this hypothesis predicts that OBLI clusters, with a /V.CCV/ preferential segmentation, should have higher syllabification consistency than other clusters with two legal onsets and /VC.CV/ preferential segmentation. For first syllable repetition the syllabification consistency order of these two categories is reversed. An examination of average syllabification consistencies for consonant clusters in these two groups shows higher consistency for OBLI clusters (average of 92% for pre-literate child subject group, and 99% for adult subjects) than for the others clusters (average of 79% for pre-literate child subject group, and 88% for adult subjects).

From these findings it appears that the comparisons of both preferential segmentation and syllabification consistency would suggest that tapping task responses have more in common with second syllable repetition, or as Content et al (in press) suggest, the process of syllable onset detection, than first syllable repetition, syllable offset detection. However, it must be noted that these conclusions are only based upon a post-hoc comparison of the two tasks, and that, whilst there are superficial commonalities between subject responses from second syllable repetition and tapping tasks there are also a number of important differences. For example, adult subject responses from the tapping task for the cluster /ks/ have greater similarity with the responses from condition 1 of Experiment 1. The zero onset responses found for this cluster, suggesting the influence of orthographic knowledge, are not found in condition 2 of Experiment 1, only condition 1. Therefore, there can be no clear and direct mapping of syllabification responses between the two tasks, a more comprehensive study would require additional information from experiments specifically designed for this purpose, a subject outside of the main concerns of this study.

Returning to the subject of the influence of orthographic knowledge in subjects' syllabification responses, it appears that the findings of Experiment 2 would suggest that orthographic knowledge plays only a minor role in literate subjects syllabification responses. With preferential segmentation agreement and a significant correlation in syllabification consistency between both pre-literate children and adult subject groups, the use of orthographic knowledge appears limited to special cases, such as the cluster /ks/. This is consistent with the results found in previous studies, where it was noted that when an effect of literacy was found it only had a temporary effect on children's syllabification performances when compared to those of adults (Gillis & De Schutter, 1996). In this experiment only a small proportion of all double intervocalic consonant clusters has been examined. It is possible that a wider examination could reveal significant differences between the syllabification behaviour of pre-literate children and adults.

Another possibility, which must be considered in the use of lexical stimuli, is that of the influence of morphology. As discussed in section 3.2, although none of the syllabification algorithms under analysis in this study takes account of morphological boundaries, the BDLEX lexicon enforces agreement between syllabic and morphological boundaries. For example, the syllable boundary markers in BDLEX normally segment the OBLI cluster /bl/ with a double consonant onset, in agreement with the preferential segmentation responses of Experiments 1 and 2. However, when a morphological boundary lies within this cluster the syllable boundary is moved to coincide with the prefix boundary, such that in the word ‘sublingual’ the cluster /bl/ is segmented with a single consonant onset.

Whilst it is feasible to conduct a post-hoc analysis of Experiment 2 to ascertain the possible influence of morphology upon syllabic segmentation an alternative solution is proposed. A new experiment would allow an examination of the influences of morphology (by comparison with the results of Experiment 2), and confirmation that the conclusions of Experiments 1 and 2 regarding preferential segmentation were valid outside of the group of stimuli used in their formulation. To this end Experiment 3 uses the same experimental design used in Experiment 2, however, non-word stimuli replace the lexical stimuli used in this experiment. This will allow an examination of the effects of morphology, by comparison with the results of Experiment 2, and a greater number of double intervocalic consonant clusters to be tested by releasing the constraints that bound the selection of lexical stimuli in Experiment 2.

3.4.2 Experiment 3 – Syllable Tapping Task using Non-word Stimuli with Adults and Pre-literate Children

3.4.2.1 Stimuli

Test items

The bi-syllabic non-word stimuli for Experiment 3 were constructed using a selection of consonant clusters from the eleven double intervocalic consonant categories used in Experiment 2. A total of 16 non-word stimuli were created for each consonant category (apart from LL and LF categories where there were 17 non-words), with each category containing between 2 and 11 different consonant clusters. A total of 82 different double intervocalic consonant clusters were used in the generation of all test stimuli. Both the CV structure and phonetic variability of the non-word stimuli used in this experiment were chosen to reflect the same distribution of forms used in Experiment 2. A list of all stimuli can be seen in Appendix D.

Stimuli were produced by 11 different monolingual female speakers from the Franche-Compté region of France. The speakers were instructed to produce the words from a randomised list of words with unchanging intonation contours, and with a pause of three seconds inserted between the production of each stimulus. Stimuli were recorded and presented to subjects over headphones or loudspeakers. The implication of this multiple speaker design is that it should help to generalise any experimental effects due to speaker variability. However, it must be noted that, to avoid confusion in the subjects, that each subject was tested with only one speaker's voice.

Distractor and Training items

A total of 38 distractor items were used in the tapping task, each contained between one and three syllables. In the construction of the non-word distractors none of the consonant clusters used in the test stimuli were used, with an equal number of items starting with vowels and consonants.

A total of 10 training items were used, selected using the same criteria as the distractor items.

3.4.2.2 Procedure

As Experiment 2.

3.4.2.3 Participants

Two subject groups were tested, pre-literate children and literate adults. In the first group there were 97 subjects (45 boys and 52 girls) with an average age of 4;5 (ranging from 4 to 5;1). All juvenile subjects were at the same level of schooling, approximately one and a half years before the tuition of reading and writing. The data of 16 additional children were discarded for the following reasons: lack of subject response (8), lack of interest or concentration (3), misunderstanding of the task and/or error rate on test items higher than 50% (5).

In the second subject group there were 113 subjects (59 men and 54 women) with an average age of 21;8 (ranging from 18 to 52). The data of 11 subjects were rejected because of high error rates (higher than 50% errors).

As in the previous experiment, all subjects were monolingual natives of Franche-Comté with no recorded auditory problems.

3.4.2.4 Results

Error rates

Of the 1566 responses collected with pre-literate subjects 12.07% were in error. For adult subjects 7.90% of the 1808 responses were erroneous. Errors included subject uncertainty, experimenter uncertainty, tapping errors (where the subject failed to tap when presented with a stimulus, or tapped more than once), and repetition errors. There was no significant effect of cluster type on error rates for either adult or pre-literate subjects.

Preferential Segmentation

The aims of Experiment 3 are twofold, confirmation that the preferential segmentation rules suggested in experiments 1 and 2 are valid over a wide range of double consonant cluster stimuli, and evaluation of the possible morphological influences on the syllabic segmentation of lexical stimuli.

Because of the number of consonant clusters tested in Experiment 3 there were insufficient observations to yield a significant preferential response in a cluster-by-cluster analysis. However, as the aims of this experiment only require a comparative analysis of subject responses with those of the previous experiment subject responses can be grouped according to the preferential segmentation findings of the previous experiments.

This results in two groups of stimuli, OBLI clusters, which should have significant preferential segmentation resulting in a double consonant syllable onset, and the remaining clusters, with a single consonant syllable onset. However, with the cluster /tl/ used in the construction of stimuli in this experiment, differences between Dell and Laporte's definitions of OBLI clusters become relevant to the analysis of results. Whilst Laporte defines OBLI clusters as one of the [p t k b d g f v] segments followed by either [l] or [r], Dell differs by excluding /tl/ and /dl/ clusters from the OBLI group. As an interim solution, putting this theoretical controversy aside for the moment, two analyses of preferential segmentation were performed, one where /tl/ was classified as part of the OBLI group, and another where it was not. Segmentation responses were analysed separately for each of the consonant cluster groups. In the OBLI group there were 13 consonant clusters (4 FL and 9 PL), with 321 observations in the pre-literate subject group, and 400 in the adult subject group. The non-OBLI group contained 68 consonant clusters, with 1041 observations in the pre-literate subject group and 1247 in the adult subject group. There were also 15 observations for the cluster /tl/ in the pre-literate subject group, and 18 in the adult subject group. These observations were added to the OBLI group for the first analysis of preferential segmentation, and the non-OBLI group for the second.

For each cluster and subject group, a χ^2 was calculated on the three possible responses, double consonant syllable onset (V.CCV), single consonant syllable onset (VC.CV), or zero onset (VCC.V), to determine whether there was a significant preferential response. In each case, and for each classification of the cluster /tl/ (OBLI or non-OBLI) the χ^2 was higher than the critical χ^2 at p=.05 (5.991 (df=2)), indicating a preferential segmentation response. This pattern of results is in agreement with those of experiments 1 and 2, showing that all clusters were segmented with a single consonant onset, apart from OBLI clusters (both definitions), which were segmented with double consonant onsets. Again, as in Experiment 2, there were no differences in preferential segmentation between pre-literate child and adult subject groups.

3.4.2.5 Discussion of Experiment 3

In Experiment 3 there was a similar pattern of preferential segmentation behaviour to that found in Experiment 2, with all clusters segmented with a single consonant onset, apart from those classified as OBLI's, that were segmented with a double consonant onset. However, the classification of the /tl/ cluster as both OBLI and non-OBLI, as in the previous section, is a little unsatisfactory. To resolve the theoretical differences regarding the /tl/ cluster it is necessary to compare the segmentation responses of this cluster with those of the non-controversial OBLI's. Subject responses for the pre-literate subject group reveal indeterminate segmentation for /tl/, with 53% of responses for /.tl/ and 47% for /t.l/. With average syllabification consistency of 78% for the other OBLI clusters, as can be seen in Figure 10, it is clear that there are considerable differences between the segmentation of /tl/ and the other OBLI's. Segmentation responses for /tl/ from the adult subject group are much clearer, with 33% of responses for /.tl/, 62% for /t.l/ and 5% for /tl./ with χ^2 at p=.05 (5.991 (df=2)), indicating a preferential segmentation response of /t.l/. These results suggest that the OBLI classifications used by Dell, where /tl/ and /dl/ are excluded from the OBLI group, is more accurate than that proposed by Laporte. This finding is also supported by the Law of Initials of the Legality Principle, out of all of the OBLI clusters defined by Laporte all are legal onsets, excepting /tl/ and /dl/, which would suggest that they should be segmented with a single consonant onset.

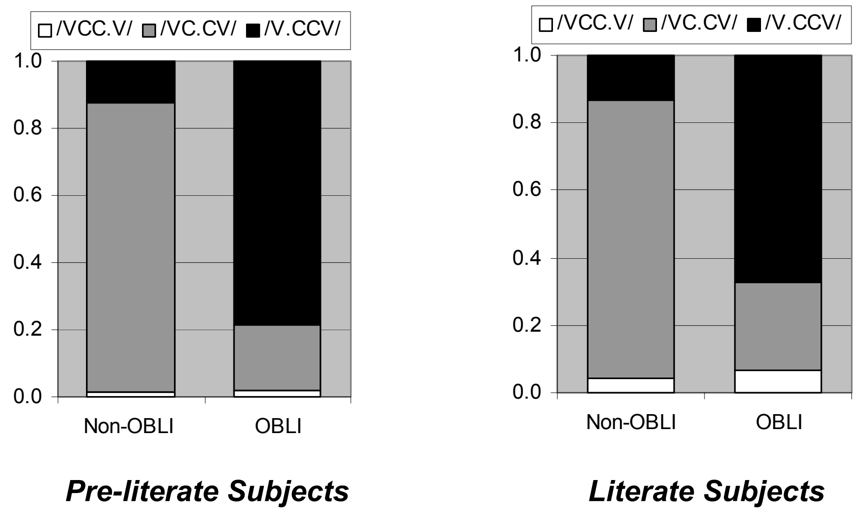


***Figure 11** Distribution of subject syllabification responses in Experiment 3 for pre-literate children and literate adults for OBLI (excluding cluster /tl/) and non-OBLI consonant clusters.*

Whilst preferential segmentation responses for the two subject groups are identical, an examination of the distribution of segmentation responses for OBLI and non-OBLI cluster groups, shown in Figure 11, shows that the differences in syllabification consistency between subject groups are the opposite to those found in Experiment 2. For child subjects there was relatively little difference, with 86% syllabification consistency for non-OBLI clusters, and 78% for OBLI clusters in Experiment 3, compared with 83% for non-OBLI clusters and 84% for OBLI clusters in Experiment 2. However, in the adult subject group syllabification consistency fell from 92% for non-OBLI clusters and 99% for OBLI clusters in Experiment 2, to 82% in non-OBLI clusters and 67% in OBLI clusters in Experiment 3. These results suggest that when faced with non-words pre-literate subjects have little difficulty in adapting to these ‘unnatural’ stimuli, however, adult subjects show greater uncertainty when it comes to segmenting non-words. These findings are also reflected in the relative increase in error levels between Experiments 2 and 3. Whilst error rates for pre-literate subjects in Experiment 3 were 53% higher than those of Experiment 2, error rates in the adult subject group increased by a factor of six (6.05 to be exact).

To summarise, examination of syllable segmentation results of Experiment 3 suggest that the preferential segmentation rules formulated in Experiment 1 are valid over the wide range of double intervocalic consonant clusters tested in Experiment 3. This allows greater confidence that these rules are generally applicable, and are not simply limited to the set of stimuli used in the previous experiments. In addition, the collection of responses from stimuli containing the /tl/ cluster, has helped to resolve the theoretical controversy surrounding the definition of the OBLI cluster, with evidence pointing towards Dell rather than Laporte's classification. Also, with preferential segmentation agreement between Experiments 2 and 3, there is no evidence to suggest that the syllabic segmentation of the words in Experiment 2 were affected by morphological boundaries. Differences in the syllabification consistency of adult subjects between the two experiments would also appear to support this conclusion. If the syllabic segmentation of the words used in Experiment 2 were affected by morphology then there would be lower syllabification consistency in this experiment than the non-words used in Experiment 3. However, the disparity between syllabification consistency in the two experiments was the reverse, with higher syllabification consistency in Experiment 2 than 3.

3.4.3 Summary

In Section 3.3.5.4 it was suggested that the distribution of subject segmentation responses for the cluster /ks/ in Experiment 1 were subject to influence due to the use of orthographic knowledge. It was thought that because this cluster is most often represented by a single grapheme, 'x', that this representation caused ambiguity as to whether the cluster contained one or two segments. However, whilst, in the case of /ks/ the influence of orthographic knowledge is clear and straightforward, it highlighted the possibility that orthographic rules may play a broader role in the segmentation responses of Experiment 1.

To investigate the extent of the influence of orthographic knowledge upon metalinguistic syllable segmentation tasks, the syllabification behaviour of two subject groups were compared in Experiment 2, those of pre-literate children, and literate adults. With knowledge of syllabic structure, but none of orthography, it is possible that the differences between the syllable segmentation behaviour of pre-literate children and literate adults could be due to the acquisition of orthographic knowledge (with all the precautions mentioned in the introduction of section 3.4).

The preferential segmentation results of Experiment 2 suggest that the role of orthographic knowledge in metalinguistic syllable segmentation tasks is limited to a few special cases. The only clear difference in the distribution of syllabification responses between the two groups was in the /ks/ cluster. In pre-literate subjects the distribution of segmentation responses was no different to that of the other PF clusters. However, the distribution of segmentation responses in adult subjects were similar to those found in condition 1 of Experiment 1, with the high incidence of zero consonant onset responses thought to typify the effects of orthographic bias upon the segmentation responses of /ks/. These findings suggest that the effects of orthographic bias upon subjects' segmentation responses are slight, limited to only a few specific clusters. Therefore the caution expressed in Section 3.3.5.4, that the preferential segmentation responses given across the whole range of stimuli are being influenced, or even decided, by 'learnt' orthographic rules, rather than unconscious segmentation, are unfounded.

However, with the use of lexical stimuli in Experiment 2, another factor comes into play which may have an effect upon syllabic segmentation; that of morphology. In the BDLEX lexicon syllable boundaries are moved from their 'normal' positions (segmented according to the nature of the consonant cluster) to coincide with morphological boundaries. To investigate the possible effect of morphology upon syllabic segmentation an additional experiment was performed. Experiment 3 used the same procedure and subject groups as Experiment 2. However, non-word stimuli replace the lexical stimuli used in the previous experiment. The use of non-word stimuli allows an examination of the effects of morphology, by comparison with the results of Experiment 2, but also allows a greater number of double intervocalic consonant clusters to be tested by releasing the constraints that bound the selection of lexical stimuli in Experiment 2. With the distribution of segmentation responses in Experiment 3 similar to those found in the previous experiment there is no evidence to suggest that morphology has a major effect on syllabic segmentation. Also, as subjects use the same pattern of syllable segmentation to that found in Experiments 1 and 2 it is more likely that the pattern of preferential segmentation found in these experiments is generally applicable, and is not limited to the set of stimuli used in the previous experiments. An additional advantage of the large test set of stimuli used in Experiment 3 is that an examination of the subject responses for the cluster /tl/ allows the theoretical differences in the definition of the OBLI cluster to be resolved. Subject responses would appear to support Dell's definition of the OBLI rather than Laporte's, as subject responses for the cluster /tl/ (which is defined as OBLI by Laporte, but not Dell) are inconsistent with the pattern of segmentation expected for an OBLI cluster.

3.5 Comparison of French-Specific Syllabification Algorithms and Syllable Perception

Taking the preferential segmentation obtained from syllable perception Experiment 1, it is possible to compare segmentation results from the syllable perception experiments with the theoretical responses of the five syllabification algorithms, plus the boundary markings given by BDLEX. Each of the five syllabification algorithms was applied to the test set of consonant clusters and singletons used in the perception experiments. Each resulting syllabification was compared to that of the preferential segmentation obtained from the perception experiment. A similar comparison was also made between the results of the perception experiment, and the syllable boundary placements used for the test set of clusters and singletons in BDLEX (those that are not affected by prefix or lexical boundaries).

In the case of the cluster /vr/, with its different, ambisyllabic, segmentations found in the first and second experimental conditions of Experiment 1 it was decided that the segmentation /.vr/ should be used as the reference for this cluster. As this was the preferential segmentation found in Condition 2 of Experiment 1 and in Experiments 2 and 3 it was thought that this segmentation would be the best representation for this cluster.

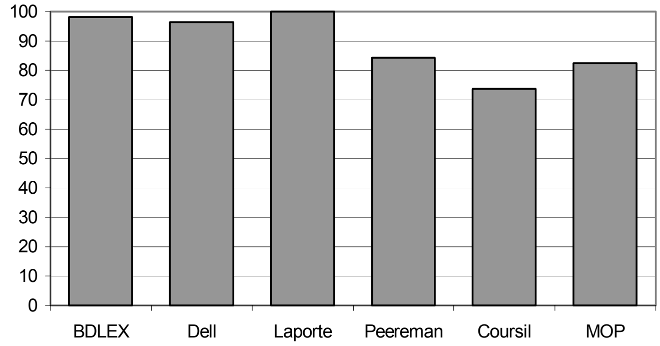


Figure 12 *Percentage agreement between the theoretical syllabification predictions experimental syllabification responses.*

As can be seen from Figure 12, the agreement between theoretical and empirical syllabification is relatively high, with a minimum agreement of 73% of consonant clusters/singletons (for Coursil). Of the five syllabification algorithms it is the two algorithms based upon the phonotactic regularities of French that show the best agreement with the subject's responses, with the algorithm of Laporte showing complete agreement with their preferential segmentation. Using this test set, the algorithm of Dell differs with that of Laporte on only one feature category, that of PFP. Whilst Laporte, and the subjects, segmented this cluster as /PF.P/, because Dell only allows a single coda, he syllabifies this cluster as /P.FP/. For the other tri-consonant feature categories, /NPL/ and /LPL/, there is agreement between the two algorithms because the /PL/ in the stimuli used forms an OBLI cluster, syllabified as /N.PL/.

3.6 General Discussion

In Section 1.4 a definition of the syllable was introduced, that of "a segment of speech that consists of a vowel, with or without one or more accompanying consonant sounds immediately preceding or following". This definition has the advantages of being clear, concise, and relatively free from controversy, however, it gives us no information on a subject of vital importance, where they begin and end.

However, a number of different theories of phonological syllabification have been put forward to help us flesh out the original definition. In brief, the main theories of syllable boundary placement that have been proposed are as follows. Firstly, the Obligatory Onset Principle, states that the onset of a syllable should contain, if possible, at least one consonant. The Legality Principle, consisting of the Laws of Initials and Finals, suggests that syllable boundaries are only possible if the resulting onset and coda are legal, that is, they can be found in word initial and final position. However, whilst this principle can tell us which syllable boundaries are possible, it states no preference in the event of multiple legal segmentations for a particular consonant cluster. The Maximum Onset Principle differs from the Legality Principle by setting the syllable boundary just before the maximum legal onset, irrelevant of the legality of the remaining coda. Other syllabification theories are based upon the sonority scale, or the inverse, that of consonanticity, with the syllable boundary placed at the trough of sonority found between syllable nuclei.

The first step in the comparison of these various principles of syllabification is their implementation in language specific forms. Five French syllabification algorithms were proposed, based on the different theories of syllabification outlined above. The first two of these, Dell and Laporte, are both based upon an analysis of phonotactic regularities according to the legality principle. Both regard a subset of OBLI clusters as indivisible, however these algorithms differ in that Dell suggests a minimal coda, whilst Laporte specifies the opposite, a minimal onset. Measures of sonority are used in two other algorithms, that of Coursil, who uses the sonority scale to define syllable boundaries at positions of double sonority plosion. The other, that of Peereman, places syllable boundaries at locations of consonanticity attack. This algorithm also gives special significance to the /s/ segment, allowing it to form part of the onset when proceeded with a plosive. The final algorithm is an implementation of the Maximum Onset Principle, using legal onsets gained from the search of a French lexicon.

One of the primary aims of this chapter was the clarification of theoretical syllabification, to take the various algorithms used in syllable boundary placement, highlight their similarities and differences, and compare them against the syllabification of naive listeners. The first step in this process was a practical examination of these algorithms, by applying them to all intervocalic consonant clusters and singletons found in the French lexicon BDLEX. This analysis uncovered a number of findings, firstly that for the majority of consonant clusters, and all singletons, there is general agreement in syllable boundary placement across all algorithms. However, this agreement is not equal across all types of consonant cluster. As the length of the consonant cluster increases, and the relative frequency of occurrence decreases, then there is greater divergence in the placement of the syllable boundary. Another finding regards the distribution of syllabification agreement between algorithms. Although some of the algorithms have significant philosophical and theoretical similarities, in practice, minor rule alterations will tend yield large differences in syllable boundary placement, especially as the consonant cluster increases in complexity. The only exception to this finding was in the comparison of the algorithms of Peereman and Coursil. Whilst the pattern of syllable boundary placement created by these algorithms was not identical, the differences between them were not statistically significant. Therefore it is suggested that in further study only one of these algorithms, both of which are based upon concepts of sonority, need be examined. However, the quandary lies in the selection of one of the two algorithms, that is, how to make a decision on the syllabification of the consonant clusters where the algorithms do not converge. Perhaps the best solution would be to choose the algorithm which best described the behaviour observed in human syllabification.

As has been seen, whilst the use of lexical statistics may help in gaining an understanding of the practical implication of the syllabification algorithms it is still necessary to obtain another, empirical, type of data to clarify syllable boundary placement. In order to obtain syllable boundary preferences from naive listeners, metalinguistic syllable repetition experiments were performed using a broad range of single, double, and triple consonant stimuli. Using subjects' preferential segmentation it was possible to form a test set of consonant clusters and singletons for comparison with theoretical predictions. These responses painted a simplistic picture of human syllabification, following that of the Obligatory Onset Principle, in that, in the majority of cases stimuli were syllabified with a single consonant onset. The only exceptions to this rule were due to OBLI clusters that were found to be tautosyllabic, forming a double consonant onset. Comparing these results with the predictions of the five syllabification algorithms it was found that only one, that of Laporte, fits the preferential segmentation of the subjects exactly. Whilst that of Dell is also in total agreement with single and double intervocalic stimuli, it diverges with longer stimuli as Dell only allows a single consonant coda. With these results there is little evidence to support the use of sonority, consonanticity, or the Maximum Onset Principle, embodied in the remaining three algorithms.

Similar results have also been found with other metalinguistic syllabification studies for English. Using a similar syllable repetition task, Treiman et al (1992) found that /s/ + stop and /s/ + sonorant clusters, whilst perfectly legal onsets, where not syllabified according to MOP, but split between coda and onset, only OBLI clusters were syllabified with a double consonant onset. Because these results run contrary to both the MOP, and that proposed by the sonority cycle (and also Peereman and Coursil), Treiman et al (1992) suggest that clusters beginning with /s/ do not form legal onsets. Although the results of the experiment conducted in this chapter concur with that of that study it is difficult to accept their conclusion. The metalinguistic syllabification experiments carried out in this study, using a broader range of consonant clusters than that of Treiman et al (1992), suggest that the syllabification of clusters with the /s/ segment were no different to that of any other non-OBLI cluster. Although the algorithms of MOP, Coursil, and Peereman failed to reproduce the subject's preferential segmentation for these clusters these algorithms also failed on a broad range of other clusters. In addition, examination of syllabification responses for experiments 1, 2 and 3 show that these clusters do not show exceptionally low syllabification consistency compared to other cluster. Therefore, it appears that the problems with these algorithms are far more serious than that suggested by Treiman et al (1992), and cannot be alleviated by granting 'special' status to situations, like for the /s/ clusters, where they fail to agree with human syllabification.

Because of problems with these stimuli, other evidence has been presented in defence of the use of sonority in human syllabification. It is suggested (Treiman & Danis, 1988) that, for single intervocalic stimuli, sonorants are more closely linked to the preceding vowel than obstruents. Whilst subject responses from the first experimental condition (offset detection) show a greater number of /VC.V/ responses for liquid consonants than plosive, no such effect was found in the second experimental condition, onset detection. One theory of particular relevance to this finding is that of the 'onset hypothesis' (Content et al, in press), which suggests that distinct operations are used for syllable onset and offset detection, with the former used for segmentation.

Whilst the syllabification consistency results of the experiment are in agreement with the tenets of the onset principle, with higher consistency for onset rather than offset detection, for single intervocalic consonants, this was not the case for double and triple consonant cluster stimuli. However, even though there is little difference in overall syllabification consistency between onset and offset detection for these stimuli, other analyses suggest that the underlying processes behind both operations may be different. These analyses concern the syllabification consistency differences that are exhibited between different consonant clusters. Some, PF for instance (63.6% consistency), show low levels of consistency, whilst others, like NF (98.5% consistency), show high consistency levels. Two possible hypotheses were suggested for this behaviour, differences in the frequency of occurrence of the consonant clusters, and the possibility of multiple legal onsets for a particular consonant cluster. Analyses showed that the number of legal onsets available for a consonant cluster held the key to predicting syllabification consistency. An examination of the results of Experiment 1 led to the suggestion of the ‘ambisyllabic onset hypothesis’, based upon the Law of Initials of the Legality Principle and theories of ambisyllabicity. This hypothesis predicts a higher proportion of increased onset ambisyllabic responses (/VC.CV/ response in Condition 1, /V.CCV/ in Condition 2) if there is a legal syllable onset before the first consonant in an intervocalic consonant cluster/singleton. This results in higher syllabification consistency for condition 2 than condition 1 when the syllable boundary of the preferential segmentation lies before the first consonant, with the reverse disparity if the preferential response lies behind the consonant. In the light of these findings it seems unlikely that syllable onset detection has special status in syllable segmentation, as the tenants of the onset hypothesis are only met in singular intervocalic consonants and a small minority of consonant clusters.

Whilst the 'ambisyllabic onset hypothesis' may allow the prediction of differences in syllabification consistency between syllable onset and offset detection, it takes no account of a number of other factors that may affect subjects' segmentation responses. Because of the nature of the task used in Experiment 1, it is possible that metalinguistic knowledge, such as that used in orthographic tasks, may influence subject responses. In Experiment 1 it was suggested that the unusual distribution of segmentation responses for the cluster /ks/, with a high proportion of zero consonant onset responses, was due to orthographic bias. This cluster is unusual in that it is most often represented by a single grapheme, 'x', therefore it was thought that this representation can cause ambiguity as to whether the cluster contains one or two segments. However, the case of /ks/ raised the possibility that orthographic rules could play significant role in the segmentation responses of Experiment 1. The effects of orthographic bias have also been noted in an examination of syllabification in singular intervocalic consonants (Treiman & Danis, 1998). In this study syllabification responses were elicited using an oral syllable reversal task, where subjects repeat a word but reverse the syllable order. The effect of orthographic bias was examined through the use of two types of stimuli, those where the singular intervocalic consonant was represented by a single grapheme (such as in 'pro*p*er'), and others that were represented by a double grapheme (such as in 'co*mm*and'). Subject responses showed that there were significantly more ambisyllabic responses when the intervocalic consonant was represented by a double rather than single grapheme. The findings of Treiman and Danis (1988), and those of Experiment 1, show that spelling may affect subject's performance in metalinguistic tasks that do not specifically require its use. However, what is not known is whether this influence is limited to a few special cases, or has a general effect on the syllabification strategies used by listeners.

To investigate the scope of orthographic bias upon the responses of experimental subjects the syllabification behaviour of two subject groups were compared in Experiment 2, those of pre-literate children, and literate adults. With knowledge of syllabic structure, but little or no knowledge of orthography, differences in the syllable segmentation behaviour of the two subject groups should highlight syllabification effects due to the acquisition of orthographic knowledge. A comparison of the segmentation responses of the two subject groups in Experiment 2 revealed that the only differences in the syllabification behaviour of these two subject groups were limited to stimuli containing the /ks/ cluster. Therefore, it appears that the influence of orthographic bias upon syllable segmentation is limited to a few special cases, such as the cluster /ks/, and has no effect on the general syllabification strategies of experimental subjects, as had been feared. Another conclusion that can be drawn from the results of Experiment 2 concerns the possibility of the contamination of segmentation results due to task specific influences. As Experiments 1 and 2 used different segmentation tasks, syllable repetition in Experiment 1 and the tapping task in Experiment 2, with similar preferential segmentation between the two, it is unlikely that the results of either task suffers from contamination due to task specific effects.

With the preferential segmentation results of Experiment 2 in agreement with those of Experiment 1, it is likely that the segmentation distinction between OBLI and non-OBLI clusters can be used as a general syllabification rule, and is not a result of orthographic bias. However, for greater confidence in these findings, and also to test the possible effect of morphology upon syllabic segmentation, an additional experiment was performed over a wider range of stimuli. Experiment 3 used the same procedure and subject groups as Experiment 2, however, in this experiment non-word stimuli replace the lexical stimuli used in the Experiment 2. The use of non-word stimuli allows, not only an examination of the effects of morphology, by comparison with the results of Experiment 2, but also a greater number of double intervocalic consonant clusters to be tested. With the results of Experiment 3 in agreement with the distribution of segmentation responses in Experiment 2 there is no evidence to suggest that morphology has a major effect on syllabic segmentation. This result also allows greater confidence that the distinction between OBLI and non-OBLI segmentation may be applied as a general syllabification rule, and is not simply limited to the test set of stimuli used in its formulation.

As for syllabic segmentation preference laws, it is suggested that those of Laporte, with syllable boundary placed before the last segment which is not a glide, with a subset of OBLI's ([p t k b d g f v] followed by [l r]) treated as a single segment, are closest to French behaviour. However, the results of Experiment 3 suggest that the definition of OBLI’s used by Laporte should be replaced by those of Dell, who excludes /tl/ and /dl/ clusters from the list of OBLI’s given by Laporte. In essence these rules form that of a Minimum Allowable Onset Principle, that the syllable boundary be placed such that the onset is of minimum length as long as it is in agreement with the Obligatory Onset Principle and the special treatment of OBLI clusters.

However, whilst a model of syllabification has been proposed which satisfies the preferential segmentation of the experimental subjects, a number of questions still remain. Firstly, will these theories reflect the syllable boundary placement of listeners outside of the test set of consonant clusters? In Experiment 3 syllabification responses were elicited over a wider range of double intervocalic consonant clusters than used in previous experiments, revealing the same pattern of preferential segmentation found in previous experiments. Whilst this allows greater confidence in Laporte's model of syllabification it must be noted that only 91 of the 431 possible consonant clusters and singletons have been analysed in the experiments presented in this chapter. Another question concerns the findings of divergent syllabification consistency levels for differing consonant clusters. Although a possible solution, namely the ambisyllabic onset hypothesis, has been proposed as to *when* low syllabification consistency occurs, no explanation can be given to describe the mechanisms that drive subjects in their choice of syllable boundaries in these cases. It must also be noted that this study of syllable segmentation was made in French, a language with a relatively clear syllabic structure. What then of other languages, such as English, in which the role of the syllable is more complex or opaque? If current syllabification models cannot explain French syllabic structure with accuracy, then it is likely that greater problems would arise in the prediction of syllable boundaries in English.

One of the reasons for syllabification inconsistencies might lie in the approaches used in current phonological syllable segmentation theory. As has been seen, the principles and algorithms that have been proposed for syllabic segmentation have only a single deciding cue for syllable boundary placement, the nature of the intervocalic consonant cluster. However, analyses have shown that this cue is insufficient in the accurate prediction of syllable boundaries, with varying levels of segmentation consistency for different consonant clusters. Therefore, it appears that improvements in syllable segmentation theory will require the examination of cues that lie outside of the consonant cluster.

4 Syllabification and the Syllable Nucleus

4.1 Introduction

As was seen in Section 2, a review of some of the most well known studies of syllabification, the theories and principles that have been proposed to solve the problem of syllabic segmentation are many and varied. Some of these theories are based upon analyses of phonotactic regularities. These suggest that there are a number of overriding principles behind the placement of the syllable boundary, such as the obligatory onset principle, or the legality principle. Other theories, such as the Principle of the Irregular Coda, go further, stating that illegality must be borne by the syllable onset rather than the offset. The preference for syllable onsets over offsets is also inherent in the Maximum Onset Principle, stating a preference for maximal legal onsets, at the cost of the legality of the syllable offset. Other approaches to the problem of syllable segmentation are based upon the sonority scale, or the inverse, consonanticity, where segments are built upon the syllable nucleus (both onset and coda) as long as they are of a lower sonority to that of the preceding segment.

Although there are considerable philosophical and practical differences amongst these theories they all share a fundamental tenet, that syllabification decisions are made on the characteristics of the intervocalic consonant cluster or singleton. Whilst syllabification cues outside of the intervocalic consonant cluster have been suggested, such as resyllabification or ambisyllabicity due to stress patterns (e.g. Kahn, 1976; Selkirk, 1982), the only universal feature of all syllables, namely the syllable nucleus, has been largely excluded from syllabification theory.

The syllable nucleus usually consists of a vowel, or according to Selkirk's (1982) hierarchical model of the syllable, may also consist of, or contain, a sonorant consonant, onto which are appended optional onset and coda consonants. With the nucleus forming such a defining and central role in the definition of the syllable, it seems unusual that it could have no influence over the definition of the syllable boundaries. However, in all of the theories of syllable segmentation discussed thus far, the nuclei of syllables are considered as little more than start and end points in the search for the syllable boundary. It is the arrangement of the consonants found between these start and end points that are used as the sole indicator of syllable boundary placement. The results of

Experiment 1 show that, for a given intervocalic consonant cluster or singleton, the preference laws proposed by Laporte (with the OBLI definition by Dell) are capable of predicting the majority of subject segmentation responses. For some intervocalic consonant clusters, such as those with a single legal onset, the Laporte's preference laws are capable of predicting up to 98% of subject responses. However, for those with multiple legal onsets, where subjects have a 'legal' choice in the placement of the syllable boundary, the percentage of subject responses in agreement with Laporte's theories can drop as low as 62%. In these cases it would appear that phonological information on the intervocalic consonant cluster/singleton is insufficient for the prediction of a substantial proportion of subject segmentation responses. This finding could indicate that subjects were using additional, uncontrolled, syllabification cues outside of the intervocalic consonant cluster/singleton when there is ambiguity about the syllabification of a particular cluster. A possible source of these additional cues might lie in the only uncontrolled segments used in the stimuli for Experiment 1, the vowels.

4.2 Vowel Duration and Syllable Structure in English

For the English language a number of studies (Barnwell, 1970; Pulgram, 1970; Umeda, 1975) have suggested that the distribution of vowels affects syllable structure, more specifically, the distribution of long and short vowels.

In English, along with certain other languages, vowels can be classified as either short, also known as lax, [ɪ,e,a,ʌ,ɒ,ʊ,ə] which require greater articulatory energy to produce than long, or tense, [iː,ɑː,ɔː,uː,ɜː] vowels. In a study of vowel duration, Umeda (1975) observed that [ɪ,e,ʌ] vowels were only to be found in closed syllables. Pulgram (1970) extends this selection of vowels by stating that all of the short vowels, excluding the schwa, are excluded from the word-final position. Such vowels are also known as *checked* vowels. Therefore, for words like 'petting' or 'potting', both with checked first vowels, the intervocalic consonant is assigned to coda of the first syllable. The remaining vowels, long vowels and schwa, may occur both syllable-medially and syllable-finally, and thus have no influence over syllable structure.

The formulation of Barnwell (1970) goes further than that suggested by Pulgram, in suggesting syllabification constraints due to long as well as short vowels. This formulation was drawn from the findings of an experiment to test the effect of syllable boundaries on vowel and consonant duration. The results of this experiment suggested that the presence of a syllable boundary between a vowel and following consonant reduces the difference in the lengthening of the vowel usually seen with a voiced following consonant. In addition, it was also found that if the subject marked the first syllable as closed there was a tendency to increase the duration of the vowel to support this segmentation. These, and other analyses, resulted in a hypothesis sometimes referred to as the *Barnwell Constraint*.

This suggests the partitioning of vowels into two classes, much like the *checked* classification of Pulgram.

Class I vowels: The set of vowels that can end an English word (long vowels and diphthongs, schwa)

Class II vowels: The set of vowels that cannot end an English word (short vowels, except schwa)

If there are multiple ways to syllabify a word, such that phonotactic and morphological constraints are satisfied, preference is given to the syllabification that assigns Class I vowels to open syllables and Class II vowels to closed syllables. An example of this behaviour can be seen in the syllabification for the word 'recall', segmented as 're.call' when produced with the Class I vowel [iː], and 'rec.all' when produced with the Class II vowel [e].

The syllabification of the previous example, 'recall', is of particular interest, highlighting the similarities between the influences of vowel length and stress on syllable structure (see Section 2, page 20, for a summary of the use of stress in syllabification). According to the obligatory onset hypothesis, the intervocalic consonant of 'recall' should be segmented as the onset of the second syllable. However, because there is stress on the first syllable, and not the second, the [k] should be resyllabified, moving from the second to the first syllable (according to Selkirk, 1982), or be ambisyllabic (according to Kahn, 1976). Similarly, the use of a class II first vowel would suggest the same segmentation.

Church (1983) compared the use of stress and vowel length as alternatives to the MOP when phonotactic constraints yielded ambiguous syllable boundary placement. This study highlighted potential problems in stress resyllabification connected with multiple medial consonants. For example, the stress patterns of words such as 'monkey' or 'bunker' would suggest that both medial consonants were contained in the first syllable. Two alternatives were suggested to circumvent this problem: 'limited stress syllabification', and 'vowel resyllabification', an implementation of the Barnwell Constraint. Limited stress syllabification replaces stress syllabification by suggesting that 'all stressed syllables must have final specifications (which may include glides or vowel elongation) (Fujimura and Lovins, 1982)'. Both of these methods allow for the multiple syllabifications for words such as 'record', whilst also solving the problem of words like 'monkey'. Both methods move the first consonant towards a lax vowel, whilst leaving the second consonant in the second syllable (fulfilling the final specification clause). Although there are no examples that suggest any practical differences between these methods, Church states that, for certain applications, information on vowel length is more readily available than that of stress, therefore vowel resyllabification might be of greater use than limited stress resyllabification.

Evidence to support the hypotheses relating vowel length and syllable aperture (the opposition between open and closed syllables) was also found in a psycholinguistic study of the syllabification of intervocalic consonants (Treiman and Danis, 1988). In this study, oral syllable reversal and written tasks were used to elicit syllable boundary decisions for bi-syllabic words with first syllable stress. Two classes of stimuli were used, the first with short first vowels, and the second with long first vowels. Analyses of results revealed that both /VC.V/ and ambisyllabic responses were more common for stimuli with short vowels than those with long vowels. While these findings would appear to support the theories of Pulgram and Barnwell, the authors of the study are sceptical. They state that although /V.CV/ responses were more frequent than /VC.V/ responses for long vowels, the results varied as a factor of the type of consonant used. The same interaction with the consonant type was also seen in segmentation responses for short vowels. They also state that the differences between short and long vowels could be an artifact of the tasks used. As subjects indicated their response by producing the first syllable of the stimulus last, it is possible that the response was coloured by production constraints. That is, when the first syllable contained a short vowel, subjects may have been more likely to close the syllable with a consonant, to produce a phonologically legal response, with the reverse case for long vowels. Also, as the subjects pronounced the stimuli in the written task, it is possible that similar effects could also colour responses in this task.

4.3 Vowel Aperture and Syllable Structure in French

Whilst there is no vowel length distinction in French, Pulgram (1970) suggests that the distinction in vowel aperture for French may play a similar role in syllabification as vowel length in English. It is suggested that there is a gradual merging of the six mid vowels (/o,ɔ,e,ɛ,ø,œ/) in such a way that they are grouped into three pairs of phonetically neighbouring open and close types (referred to as the three *vowel types*, /o//ɔ/, /e//ɛ/, and /ø//œ/), with open vowels (/ɔ, ɛ, œ/) regularly occurring in non-final, and close ones (/o, e, ø/) in final position. It is also suggested that if this trend persists, that the aperture differentiation between the three mid vowel types will disappear, replaced with three mid vowels, each with an open and close allophone in complementary distribution with respect to the syllable.

The general trend for open vowels in closed syllables, and vice versa, suggested by Pulgram, is also evident in Tranel's (1988) summary of current French vowel production in word-final syllables, shown in Table 8. Although there is aperture neutralisation in certain cases, in agreement with the theories of Pulgram, vowel aperture still plays a contrastive role in certain situations. However, even when both open and close vowels are used in a certain syllabic context, the occurrence of a vowel whose aperture is in agreement with that of the syllable is usually a special case. In general, the distribution of mid vowels amongst open and closed syllables support the hypothesis that open vowels are generally found in closed syllables, and close vowels in open syllables.

	Open Syllable	**Closed Syllable**
/e//ɛ/	/e/ and /ɛ/ *(et and haie)*	/ɛ/ *(air)*
/ø//œ/	/ø/ generally *(eux)* /œ/ in grammatical words: *je, de, que, etc.*	/ø/ before /z/ *(-euse)* /œ/ before /r/ *(heure)* and elsewhere in general.
/o//ɔ/	/o/ - *(eau)*	/o/ before /z/ *(ose)* /ɔ/ before /r/ *(or)* elsewhere generally: /o/ - *au, ô (haute, hôte)* /ɔ/ - *o (hotte)*

Table 8 *Distribution of mid vowels in the production of word-final syllables* (from Tranel (1988))

Although there seems little doubt that there is an inverse relationship between vowel and syllable aperture in French, a question of vital importance, at least for theories of syllabification, is the direction of this relationship. That is, do open syllables encourage the production of close vowels, or is it that close vowels encourage open syllables. Pulgram (1970) suggest that it is the latter case, stating:

'*Syllable-finally is also section-finally in terms of phonotactics, and therefore once it is established that only close vowels occur section-finally and only open ones section-medially, the occurrence of one or the other allophone implies the locus of the syllable boundary, whose terminal behaves phonetically like that of a section boundary. It is therefore not the syllable boundary itself which signals a meaningful distinction, nor of course the allophone, but the occurrence of an* *allophone that is regularly, being either of the section-final or section-medial variety, either followed or not followed by the phoneme of pause.'*

Extending this case to a multi-syllabic context, a change in the aperture of a vowel would induce a change in the syllable structure of the carrier word. Pulgram (1970) uses the example of the word 'peureux' which may be /pø.rø/ phonotactically, but also /pœr.ø/ in analogy of 'peur' (/pœr/) from which it is derived.

As yet, no psycholinguistic studies have examined the possible relation between vowel aperture and syllabic structure, even though the conclusions of Pulgram (1970) and the findings of Treiman and Danis (1988) might suggest that aperture, like vowel length in English, could induce changes in syllabic structure. Although no mid-vowels were used in the stimuli of Experiment 1, it is possible that changes in the aperture of the vowels used in the stimuli could have influenced subject's segmentation decisions. As segmentation responses in this experiment were grouped according to their consonant cluster or singleton, without taking changes in first vowel aperture into account, this behaviour would have manifested itself as syllabification inconsistency. Therefore, in light of the syllabification consistency findings of Experiment 1, it was decided that the theoretical relation between vowel aperture and syllable structure merited further investigation.

To investigate the hypothesis that close vowels induce open syllables, and open vowels induce closed syllables, referred to as the *vowel aperture segmentation hypothesis,* a number of experiments were performed. In the first set of experiments, presented in Chapter 5, both the production and perception of vowel aperture were tested. These experiments can be considered as preparatory, designed to verify that subjects were capable of producing vowels of both apertures in a given syllabic context, and perceiving the differences in vowel aperture. The second set of experiments, presented in Chapter 6, was designed to investigate the vowel aperture hypothesis of segmentation, using two experimental paradigms. The first experiment consisted of a syllable repetition task, as used in Experiment 1, with stimuli designed to test the differences in syllable segmentation due to changes in the aperture of the first vowel. In the second experiment the same stimuli were used in a fragment detection task, an 'online' task, to clarify the results of the previous, metalinguistic, task.

5 Vowel Aperture Production and Perception

5.1 Introduction

In Section 4.3, it has been suggested that vowel aperture at the nucleus of a syllable might have an influence on the configuration of its coda. This hypothesis states that if the vowel were open, then this would encourage a closed syllable, and if the vowel were close, would encourage an open syllable. This hypothesis rests upon two assumptions, that subjects are capable of producing both open and close variants of mid vowels in non-final syllables, and that they are also capable of perceiving the difference between these variants.

A summary of the production of mid vowels in non-final syllables by Tranel (1988) reveals that, although there is a tendency to pronounce /e, œ,ɔ/ rather than /ɛ,ø,o/, there are relatively few restrictions. Most of these restrictions are a result of syllable structure. For instance, if a non-final syllable is closed by /r/, then this forces the occurrence of open vowels. The presence of /z/ is also supposed to favour the occurrence of /ø/or /o/ rather than their open counterparts. The distribution of mid vowels also depends on the degree of the aperture of the vowel in the following syllable, known as *vowel harmony*. When the following vowel is close or half-close, the occurrence of /e/ is reinforced. The rounded mid vowels /œ/ and /ɔ/ will also close to /ø/ or /o/ when followed by their close variants. In all remaining cases either open or close mid vowels may be produced in non-final syllables, allowing the variation required for vowel aperture to have a significant impact upon syllable segmentation. As for vowel perception, the very existence of open and close mid vowel phonemes signifies that subjects should be capable of differentiating between vowels of different aperture in minimal pairs.

Whilst there appears to be evidence to support both the production and perception of vowel aperture, in general, there are two potential problems that must be confronted. Firstly, it is possible that there could be differences in the use of vowel aperture, either in production or perception, due to the regional accent of the experimental subjects (most of which are Swiss French). The second potential problem arises from a subject discussed in Section 4.3, concerning the direction of the relationship between vowel aperture and syllable structure. Although Pulgram

(1970) concludes that it is vowel aperture that influences syllable structure, and not the other way around, there is no empirical evidence to support this view (other than that of Treiman and Danis (1988) for English vowel length). If the syllable structure of an utterance was decided from the nature of the intervocalic consonant cluster or singleton, and, in turn, the perception or production of vowel aperture was dictated by the syllable structure (open syllables inducing the perception or production of close vowels, and vice versa), then vowel aperture is of little use in syllable segmentation. In this case, open vowels would always be coupled with consonant clusters and singletons that would normally be segmented with a closed first syllable, and vice versa. Therefore potential syllabification cues from the vowel aperture would only serve to reinforce the syllabification cues of the intervocalic consonant cluster/singleton, rather than being an additional, independent, cue.

Because of these potential problems, two experiments are suggested to verify the assumptions of the vowel aperture segmentation hypothesis. The first experiment tests the subject's ability to produce both open and close vowels. A second experiment tests the subject's ability to differentiate between open and close vowels. It must be noted that these experiments were designed purely to verify the basic assumptions of the original hypothesis, and were not designed to offer any wider conclusions on the subjects of either vowel production or perception. To keep any dialectal behaviour constant between experiments, subjects for these experiments were drawn from a pool of 36 potential subjects, the same pool of subjects that were also used in the syllable perception and fragment detection tasks presented in Chapter 6. By using the same pool of subjects in all experiments, it is possible to be sure that, if the basic assumptions of the original hypothesis are upheld in the vowel production and perception experiments, they will also be upheld in the subsequent syllable perception experiments.

5.2 Experiment 4 – Vowel Aperture Production

This experiment examines vowel aperture production, specifically, the aperture of the vowel preceding an intervocalic consonant. If vowel aperture production is inversely related to the aperture of the syllable, as determined by the nature of the intervocalic consonant cluster/singleton that follows it, then subjects should have difficulty in producing open vowels in open syllables, and close vowels in closed syllables.

To test this hypothesis, subjects were asked to read aloud from a list of bi-syllabic words containing an intervocalic consonant. Two sets of words were prepared for the subjects. In the first set, the orthographic representation of the first vowel is supposed to favour either an open, or close vowel (Tranel, 1988), as shown in Table 9. It must be noted that orthographic cues for open or close vowels are usually limited to formalised elicitation, where pronunciation is more deliberate. For spontaneous, or rapid unmonitored, speech, vowel aperture neutralisation can take place. For the mid vowel aperture pair /e//ɛ/ the vowel /e/ will tend to be favoured, /œ/ will be favoured for the /ø//œ/ pair, and /ɔ/ for /o//ɔ/. It is expected that the production of words in this experiment will favour the use of orthographic cues, as subjects are reading aloud from a written list of words, and are also asked to take care in their pronunciation. In the second set of words the orthographic representation of the vowel is neutral. In these cases, the expected production of open and close vowels were referenced from their phonetic transcription in the BRULEX lexicon.

Vowel	**Orthographic Representation**
/e/	é
/ɛ/	è, ê, ai, ei, e
/o/	ô, au, eau
/ɔ/	o

***Table 9** Orthographic representations of mid vowels*

The usual syllabification of single intervocalic consonants, referenced from Experiment 1, results in an open first syllable. Therefore, if the location of the syllable boundary influences vowel aperture production, then subjects should show a preference for close, rather than open first vowels.

5.2.1 Stimuli

5.2.1.1 Test items

To find words suitable for use in Experiment 4, a search was made of the BRULEX lexicon for all double syllable words with a mid-vowel (/o,ɔ,e,ɛ,ø,œ/) in the first vowel position followed by any of the /t,d,v,s,n,m,r,l/ intervocalic consonants. Words were selected from those found in this search for three stimuli groups. The first group consisted of 26 words with an orthographic representation that indicated the production of an open first vowel. The second group consisted of 18 words with an orthographic representation that indicated the production of a close first vowel. The third group consisted of 9 words where there was no orthographic indication of first vowel aperture. For the latter group the phonetic labelling found in BRULEX for five of the words indicated that the first vowel was produced with a close vowel, for the remaining words the first vowel was labelled as an open vowel.

A full list of stimuli can be seen in Appendix E.

5.2.1.2 Distractors

An additional 25 bi-syllabic words with double intervocalic consonants clusters were added to the list of stimuli to act as distractors. These words had the same distribution of mid-vowels in the first vowel position as those of the test items, but different consonant clusters from those used in the test items.

5.2.2 Procedure

Before the experiment, subjects were asked to examine the randomised list of test and distractor words and to indicate any unknown words, or those where the pronunciation was unclear. For these items, subjects were told to produce the word as naturally as possible, using similar sounding words (of their choice) as a guide to their pronunciation. During this phase, the experimenter was careful not to repeat stimuli items or pronunciation examples, nor was any indication given about the desirability of the pronunciation chosen by the subject. When the subjects were confident in the pronunciation of the words, they were placed in an acoustic booth and asked to read aloud from the list of stimuli, pausing for at least two seconds between each word. All words were captured using a Sennheiser ME80 microphone and Rane MS1 microphone pre-amplifier onto a Sony ZA5ES DAT recorder.

5.2.3 Participants

All 36 participants (34 female, 2 male) were students of the Université de Genève, with an average age of 22;1 (ranging from 18 to 39), and were native speakers of French with no known hearing defects. They received course credits for their participation.

5.2.4 Transcription

The target vowel in the words produced by the subjects was transcribed by two trained phoneticians, neither of which had any knowledge of the aims of the experiment. Both transcribers were monolingual native speakers of French, chosen to represent two regional French accents, the first from Paris, and the second from Geneva, representing the same regional accent as the vast majority of experimental subjects. The phoneticians were presented with the orthographic representation of the words produced by the experimental subjects (in the same order as that produced by the relevant subject), and asked to make a broad phonemic transcription of the vowel highlighted for each word (for example, in the word ‘s*eu*lement' the first vowel is that which requires transcription). The phoneticians were presented with the words over headphones directly from the DAT tape to which they were recorded.

5.2.5 Results

5.2.5.1 Transcription Validity & Production Errors

To test the transcription consistency of the results a comparison was made of the broad phonemic representations of the vowels transcribed by the two phoneticians. This analysis revealed that 1818 of the 1909 vowel productions (95.29%) were transcribed with the same broad phonemic representation by both phoneticians. To examine the similarity of these transcriptions the data from both phoneticians were cross-tabulated, taking into account all vowels transcribed by either phonetician (all mid vowels, plus /a/) and compared with a McNemar test. This test revealed that there was no significant difference between the two transcriptions of the vowels as $p > 0.05$ ($p=0.207$).

Of the 90 disputed vowel transcriptions there was a significantly greater number of inconsistencies due to differences in *only* vowel aperture (for example, one transcription of the vowel as /o/ the other as / ɔ/), than those involving differences due to combinations of any of the other factors, such as backness etc. With 73 inconsistencies in the former category, and 13 in the latter a test of the goodness of fit revealed a significant difference between the distributions of errors in these categories ($\chi^2 = 33.62$, $df = 1$, $p < 0.05$).

As all transcription disparities falling into the former category were found with mid vowels (/o,ɔ,e,ɛ,ø,œ/), it was possible that that differences in the perception of certain of the vowel types by the transcribers could be the cause of the relatively high proportion of vowel aperture transcription disparities. However, a χ^2 analysis of the distribution of the inconsistencies revealed that there was no significant difference in inconsistencies between the three vowel types (/o,ɔ/,/e,ɛ/ and /ø,œ/) ($\chi^2 = 0.12$, $df = 2$, $p = 0.94$).

Analysing the distribution of all transcription inconsistencies showed that there was no significant effect of either subject ($\chi^2 = 2.279865$, df = 51, p =1.0) or word ($\chi^2 = 2.52$, df = 35, p = 1.0) on the distribution of these discrepancies. Therefore, apart from the relatively high disagreement between transcribers concerning vowel aperture, the distribution of transcription inconsistencies appears to be evenly distributed amongst the experimental data.

Another factor that needs to be taken into account when interpreting the transcriptions of the data is phonetic expectation. If the transcribers are aware of the meaning of the words to be transcribed, as is the case in this experiment, then there is a tendency for transcribers to hear what they expect to hear, not that spoken by the subject (Oller & Eilers, 1975). Naturally, the impact of phonetic expectation is difficult to estimate, and, according to Oller and Eilers (1975) there are no methods for eliminating this effect from the transcription of utterances whose representation is already known.

Although there are no direct means available for the assessment of either the degree or distribution of this effect it is possible that the differences in the regional accent of the two transcribers could highlight this effect. The differences between the Parisien accent and that used in Suisse Romande of relevance to this experiment concern the production of vowel aperture in mid vowels. For the pairs /o,ɔ/ and /ø,œ/, Parisians will most often assign a closed vowel to an opened syllable and vice-versa (*vélo* (bike) will be pronounced /velo/, and jeune (young) will be produced as /ʒœn/). However, in the dialect used around Geneva this rule can be reversed, *vélo* will be produced /velɔ/, and *jeune* /ʒøn/, but this reversal is not systematic. As there are differences in the use of vowel aperture between the two transcribers then the effects of phonetic expectation should be evident in vowel aperture transcription disagreements. It is suggested that this effect could explain the relatively high incidence of vowel aperture disagreements between the two transcribers.

Therefore, in an attempt to minimise the effects of phonetic expectation on the experimental data used in the analysis of this experiment, data that was subject to transcription disagreement were removed from analysis. After removal of these items, 1818 remain, of which a further six were removed due subject error (4 instances where the word was not produced by the subjects, and 2 instances where the subject produced the wrong word) and another that was produced, and transcribed by both phoneticians, with the wrong vowel (one subject produced the vowel /a/ instead of either /o/ or /ɔ/ in the word ‘aussi’. This left a total of 1811 vowel productions for analyses, 95.3% of that originally intended.

5.2.5.2 Vowel Aperture Production

Words with Orthographic Cues to First Vowel Aperture

Analyses of phonetic transcriptions of the spoken words reveal that, in most cases, the production of the aperture of the first vowel was in agreement with its orthographic representation. For words where the orthographic representation suggested the use of an open vowel, 88.54% of the target vowels produced by the subjects were open, and where the orthographic representation suggested the use of a close vowel, 81.33% were close. This behaviour was common across almost all the words tested, with the majority of the vowels produced for a particular word in agreement with its orthographic representation. The only exception to this finding was for the word 'chéri', whose orthographic representation would suggest the use of a close vowel. For this word, 61% of the vowels produced by the subjects were open.

To test the significance of the general agreement between the production of vowel aperture and its orthographic representation, a test of the goodness of fit was made on the number of open and close vowel productions for the two classes of stimuli (open and close vowel orthographic representations). This analysis showed that the distribution of open and close vowel productions for both categories was statistically significant (at $p<0.05$) with a χ^2 of above 3.841 (df=1).

These results suggest that vowels of both open and close aperture could precede single intervocalic consonants, even though these consonants are usually segmented with an open first syllable. Another aspect of these results of potential interest is the difference in the consistency of vowel aperture production between words with orthographic cues for open and close first vowels. As has been seen, the proportion of vowel aperture productions in agreement with the orthographic cues of the words, referred to as vowel aperture production consistency, is higher (88.54%) for open vowels than close vowels (81.33%). To test the significance of this finding, a test of correlation was made between the consistency of vowel aperture production (the proportion of subject productions with the same vowel aperture as that suggested by the orthographic representation of the word) for each word and the aperture of the first vowel suggested by it's orthographic representation (open vowel aperture represented as 1, close aperture as 0). This test revealed that there was no significant ($r= 0.237$, $t(42)= 1.584$, $p>0.05$) correlation between these factors, indicating that the production of open vowels is not significantly more consistent than that of close vowels.

Words without Orthographic Cues for Vowel Aperture Production

In the previous section it was found that, for words with orthographic cues indicating target vowel aperture, both open and close vowels were produced in words where segmentation usually results in an open syllable. In this section, vowel aperture production is examined for words where the orthographic representation of the first vowel does not give any indication of aperture preference.

The production of first vowel aperture was analysed for each of the nine words in this stimuli group. The distribution of open and close vowel production for each word, shown in Figure 13, was statistically significant (at p=0.05) for all but one of the words, with a χ^2 of above 3.841 (df=1). For six of the nine words a significant majority of subjects produced a close first vowel, for the word 'gueulante' the production of vowel aperture was ambiguous, and for the words 'pleureur' and 'breuvage' a significant majority of subjects produced an open first vowel.

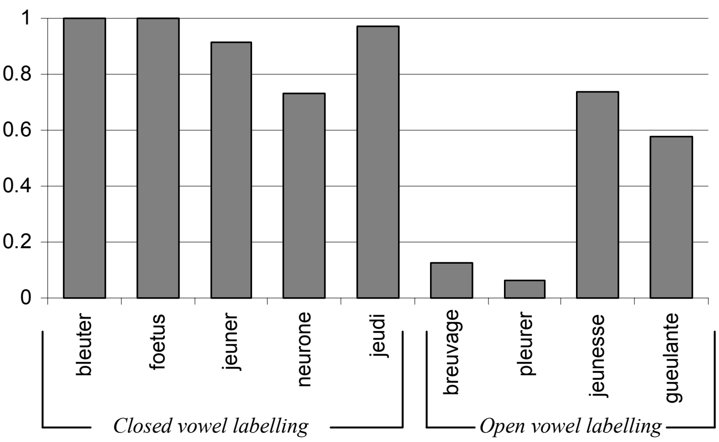


Figure 13 *Proportion of close first vowels in the production of words with no orthographic indication for first vowel aperture in Experiment 4*

These results suggest that, even without orthographic cues for vowel aperture production, there are still differences in the production of vowel aperture between words. However, it was found that the subjects did not always agree with the aperture labelling indicated by BRULEX. Even though four of the words tested were labelled with an open vowel in the BRULEX lexicon, one of these was produced with a significant proportion of close first vowels, for another the production of vowels was ambiguous. Although this finding would appear to support the theoretical link between open syllables and close vowels, the test set of words was not large enough to support any general conclusions of this nature. It is possible that, as discussed previously, vowel aperture production for the words 'jeunesse' and 'gueulante' in Suisse Romande is different to that used in Paris, the 'standard' French used for the compilation of the BRULEX lexicon. Another possibility is that the vowel aperture labelling in BRULEX is inaccurate, and that the words chosen for this stimuli group do not represent the true distribution of open and close first vowels in words without orthographic cues to first vowel aperture.

5.2.6 Conclusion

The aim of this experiment was to discover whether vowel aperture production is inversely related to the aperture of the syllable, as decided by the nature of the intervocalic consonant cluster/singleton that follows it. If this were the case, then subjects should have had difficulties in producing open vowels with the words used in this experiment, as they all contained a single intervocalic consonant, usually segmented with an open first syllable.

Analyses of vowel aperture production revealed that, for stimuli with orthographic cues to vowel aperture production, subjects had little difficulty in producing either open or close vowels, consistent with the orthographic representation of the first vowel. For words with an orthographic representation favouring an open first vowel a significant proportion of vowel productions were open, with 89% of vowel production open. Likewise, for vowels in words with orthographic representation favouring a close first vowel 81% of products were close, a significant proportion of all vowel production for these words.

Words with no orthographic cues to vowel aperture production were balanced using the labelling of vowel aperture found in the BRULEX lexicon, with five words labelled with a close first vowel, and four with an open first vowel. Analyses of vowel aperture production for these stimuli were less conclusive. Whilst there was significant agreement with the labelling of the five close vowels there was only a significantly greater incidence of open, rather than close, productions for two of the four words labelled with open vowels. Whilst it is possible that the labelling of vowel aperture for the remaining two words was incorrect it is more likely that this result was due to dialectal differences between the experimental subjects and the Parisien dialect used in BRULEX.

In conclusion, it appears that there is evidence to support to confirm the hypothesis of this experiment, that vowel aperture production is not limited by syllabic structure, as subjects were able to produce both open and close vowels in situations where segmentation would normally result in an open syllable. Although it is possible that there was, unavoidable, contamination of results due to phonetic expectation in the transcribers, it is suggested that the use of transcribers with different dialects should minimise this effect, by analysing only those items that were transcribed with the same broad phonemic category by both transcribers.

5.3 Experiment 5 – Vowel Aperture Differentiation

If the perception of vowel aperture is related to the aperture of the syllable, subjects should have great difficulty in differentiating between open and close vowels in two stimuli with matching consonantal structure (open or closed syllable). To investigate vowel aperture perception, a vowel differentiation experiment was performed. In this experiment, pairs of monosyllabic stimuli are presented to the subjects, who have to decide whether the stimuli are the same or different. Each stimuli pair has matching consonantal structure, of either CV or CVC form (to test vowel perception in both open and closed syllables), the only factor that may be used to differentiate the two stimuli is the nature of the vowel. Two sets of stimuli are used in this experiment. The first set of stimuli pairs use the phonetically neighbouring open and close mid vowels (/o//ɔ/, /e//ɛ/, and /ø//œ/) to test vowel differentiation, where the only distinguishing feature between vowels is that of aperture.

In the second set of stimuli, vowel differentiation is tested with a wide variety of vowel types to act as a kind of 'benchmark' for the previous set of stimuli pairs. The subject responses from this set of stimuli may then be compared to those which differ only in vowel aperture to discover whether these vowels are as easy to distinguish as those differentiated by a range of factors, such as nasalisation, backness and roundness.

5.3.1 Stimuli

5.3.1.1 Test items

To test the ability to differentiate between different vowels, subjects were presented with pairs of stimuli. Each of the stimuli in the pair shared the same consonantal structure, the only variable which could be used to differentiate them was the nature of the vowel. For some stimuli pairs there was agreement in the vowel used, making the two stimuli in the pair identical. However, for others, the vowel used in the first and second stimuli differed.

Stimuli pairs were organised into two categories. The first category of stimuli tested the subjects ability to differentiate between the aperture pairs of mid vowels (/o//ɔ/, /e//ɛ/, and /ø//œ/). The second category tested subjects' ability to differentiate between vowels that differed by factors such as roundness, backness, nasalisation, etc. The stimuli used in this category used a selection of the vowels [/u/, /i/, /y/, /ɔ̃/, /ɛ̃/,/o/, /ɔ/, /e/, /ɛ/, /ø/, /œ/].

In the first category (the vowel aperture category), CV and CVC stimuli were organised into pairs as described in Figure 14. For each of the three vowel types (/o//ɔ/, /e//ɛ/, and /ø//œ/)

there were four token pairs with differing vowel aperture, and six with identical aperture. This resulted in 30 token pairs for CV stimuli and 30 token pairs for CVC stimuli.

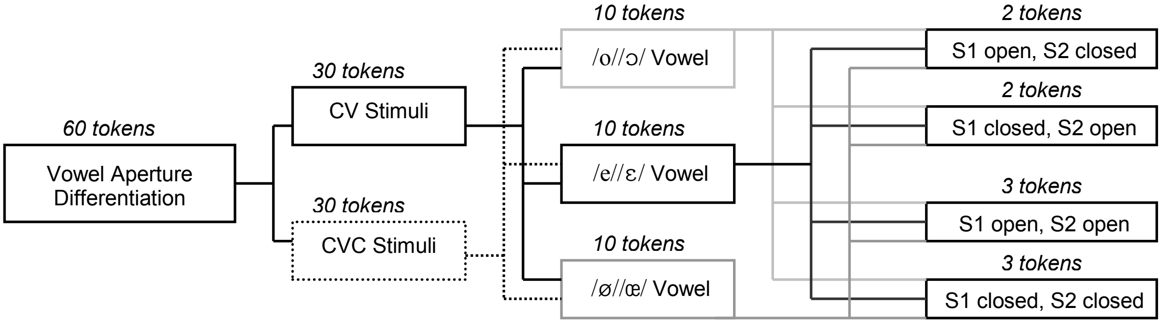


***Figure 14** Organisation of stimuli to test the differentiation of vowel aperture in Experiment 5 (S1 = first stimulus in pair, S2 = second stimulus in pair)*

In the second category (the vowel type category), CV and CVC stimuli were organised into pairs as described in Figure 15. This organisation resulted in five identical CV and CVC stimuli, one for each of the [/u/, /i/, /y/, /ɔ̃/, /ɛ̃/] vowels, and twelve stimuli pairs with different vowels. Vowels in the latter stimuli were chosen such that they differed by factors other than aperture, such as backness, roundness, nasalisation, or a combination of these, and other factors.

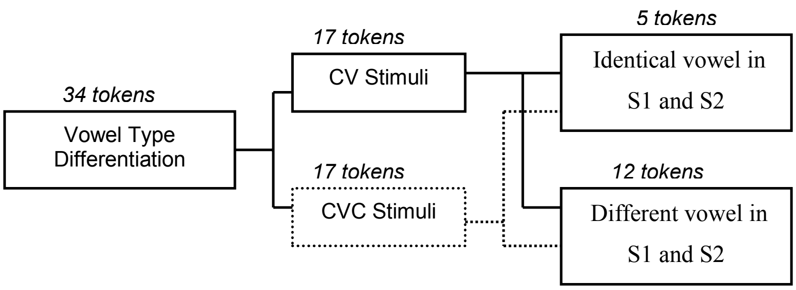


***Figure 15** Organisation of stimuli to test the differentiation of vowel type in Experiment 5 (S1 = first stimulus in pair, S2 = second stimulus in pair)*

The organisation of stimuli in both the vowel aperture and vowel type categories resulted in 46 identical stimuli pairs and 48 stimuli pairs with vowel disparity. Whilst the consonantal structure was identical for stimuli pairs, a different selection of consonants were used for each stimuli pair to increase stimulus variability. In addition, the combination of consonantal structure and vowel were arranged to ensure that all stimuli were non-lexical.

5.3.1.2 <u>Training Items</u>

A total of 20 training stimuli pairs were generated, chosen to represent a similar distribution of stimuli to that of the test items. These consisted of 10 identical pairs, 6 pairs which differed by vowel aperture, and 4 which differed by vowel type, each of which were split between CV and CVC stimuli. Training stimuli used different combinations of consonant structure and vowel to those found in the test items.

All stimuli were produced by a female adult monolingual, naïve (that is, they had no knowledge of the aims or procedure behind the experiment), native French speaker (Parisien accent) from a randomised list of non-words represented phonetically using the IPA alphabet, with a pause of three seconds inserted between the production of each stimulus. These non-words were recorded using a Sennheiser ME80 microphone and Rane MS1 microphone pre-amplifier onto Sony ZA5ES DAT recorder. Afterwards the stimuli were low-pass filtered at 11kHz and then digitised with a fidelity of 22kHz with 16 bits per sample.

5.3.2 Procedure

The experiment was performed automatically using the Psyscope experimental software running on a Macintosh computer with accompanying button box. Audio stimuli were reproduced using this software through an audiomedia sound card and presented to subjects using Beyer Dynamic DT100 headphones. During the experiment subjects were seated in an amplisilence acoustic booth.

Subjects were asked to differentiate stimuli pairs, one following the other, the second starting 300ms after the end of the first. After hearing both stimuli, the subjects were to respond by pressing one of two buttons as quickly as possible. One button indicated that the stimuli were the same, the other that the stimuli were different. Stimuli pairs were presented continuously, in random order, one every two seconds, without pause, with a short warning tone given before each stimuli pair. At the start of the experiment, a short training block of 20 stimuli pairs was presented.

5.3.3 Participants

All 19 participants (18 female, 1 male) were students of the Université de Genève, with an average age of 23;5 (ranging from 19 to 39), and were native speakers of French with no known hearing defects. They received course credits for their participation.

5.3.4 Results

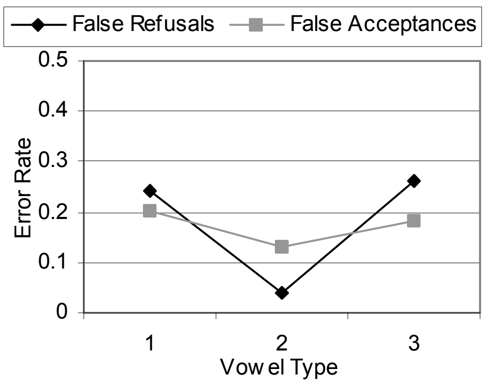
5.3.4.1 Experimental Error Rates

The only experimental errors found in Experiment 5 were null response errors, which accounted for 1.58% of the possible responses from the vowel aperture category of stimuli pairs, and 2.16% of the possible responses from the vowel type category of stimuli pairs. Errors were randomly distributed amongst the stimuli pairs in both categories.

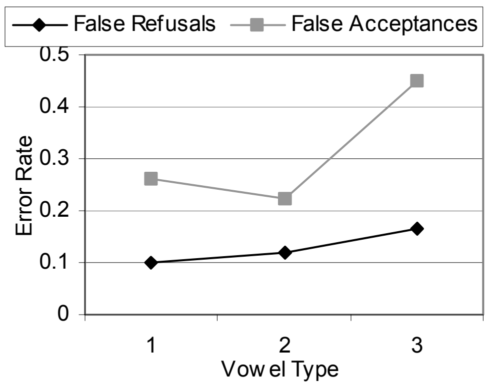
5.3.4.2 Vowel Aperture Differentiation

Analysis of all subject responses for vowel aperture differentiation stimuli pairs revealed that 81.37% of responses were correct, that is, subjects correctly identified identical stimuli as the being the same, and those differentiated by vowel aperture as different. To test the significance of this result, a test of goodness of fit was made on the distribution of correct and incorrect responses, showing that the proportion of correct responses was statistically significant (at $p<0.05$), with a χ^2 of above 3.841 (df=1).

The error rate of 18.63%, calculated over all subject responses may be broken down into two types of error. False acceptances, where the subjects responded that the stimuli pairs were the same when they were different, and false refusals, where subjects responded that stimuli pairs were different when they were identical. False acceptances accounted for 24.34% of subject responses for different stimuli pairs, and false refusals for 14.78% of subject responses for identical stimuli pairs.



CV Stimuli Pairs



CVC Stimuli Pairs

Figure 16 *False Acceptance and False Refusal error rates in vowel aperture differentiation for CV and CVC stimuli across vowel type (1 - /ø//œ/, 2 - /e//ɛ/, 3 - /o//ɔ/) in Experiment 5*

To test the possible significance of the variability in false acceptance and refusal errors over the experimental factors a log linear analysis were conducted on the frequency of erroneous responses summarised in Figure 16. Three within subject factors were analysed, consonantal structure (CV or CVC), vowel type (three types, /ø//œ/,/o//ɔ/,/e//ɛ/), and error type (false acceptance and false refusal). Because the data collected in the experiment was unbalanced, with respect to the number of stimuli with the same (36 stimuli) and different (24 stimuli) vowel aperture (see Figure 14), it was necessary to perform a post-hoc balance of the data used in this analysis. Therefore a random selection of one third of the stimuli used in the experiment was removed from the analysis such that there was an even number of observations across all experimental variables (4 stimuli, 76 observations, for each modality of the three experimental variables). Missing data was replaced with the average value of the subject responses for each particular stimulus.

The analysis revealed significant interactions between error type and consonantal structure (G^2=7.68, df=1, p<0.05), and vowel type and consonantal structure (G^2=7.98, df=2, p<0.05). Further analyses revealed that the former interaction, that between error type and consonantal structure, shown in Figure 17, described a significant difference in the distribution of false acceptance and refusal errors in CVC stimuli ($\chi^2_{df(1)}$ = 15.1, p<0.05), that did not occur in CV stimuli ($\chi^2_{df(1)}$ = 0, p>0.05).

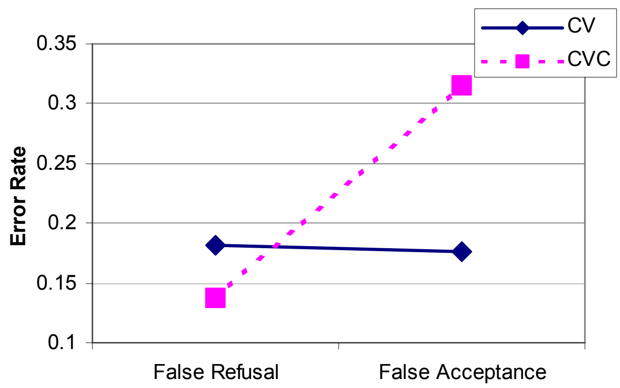


***Figure 17** False Acceptance and False Refusal Vowel Aperture Differentiation Error rates for CV and CVC stimuli pairs in Experiment 5*

A more detailed examination of the interaction between vowel type and consonantal structure, shown in Figure 18, revealed that there was an significant difference in the distribution of errors across vowel type for both CV ($\chi^2_{df(2)}$ = 10.3, p<0.05) and CVC stimuli ($\chi^2_{df(2)}$ =12.76, p<0.05). However, a comparison of error frequency for each vowel type revealed that there was only a significant difference in the error rates between CV and CVC stimuli with stimuli using the

/o//ɔ/ vowel type ($\chi^2_{df(1)} = 4.94$, $p<0.05$), and not for vowel types /ø//œ/ ($\chi^2_{df(1)} = 1.3$, $p>0.05$) and /e//ɛ/ ($\chi^2_{df(1)} = 2.56$, $p>0.05$).

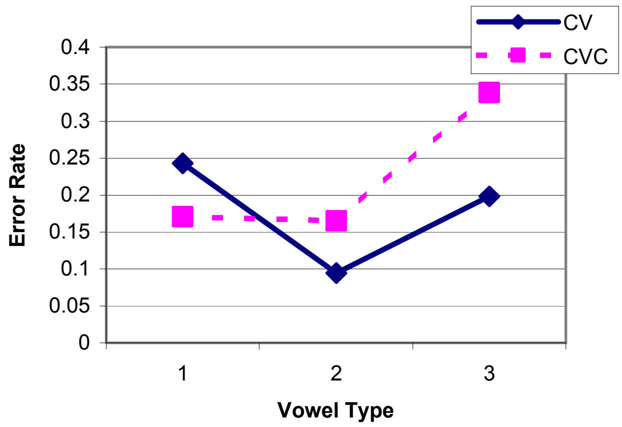


Figure 18 *Vowel Aperture Differentiation Error rates for across three vowel types (1=/ø//œ/, 2=/e//ɛ/, 3=/o//ɔ/) for CV and CVC stimuli pairs in Experiment 5*

5.3.4.3 Vowel Type Differentiation

To produce a benchmark for general vowel differentiation, for comparison with the results of the previous section, a test was made of the subject's ability to differentiate between a wide variety of vowels, differing in factors such as nasalisation, roundness, and backness.

Analysis of subject responses for these stimuli pairs revealed that 82.44% were correct, with a test of goodness of fit made on the distribution of correct and incorrect responses, showing that the proportion of correct responses was statistically significant (at $p=0.05$), with a χ^2 of above 3.841 (df=1). False acceptances accounted for 20.58% of subject responses for identical stimuli pairs and false refusals for 10.27% of subject responses for different stimuli pairs. To test the possible significance of the variability in false acceptance and refusal errors between CV and CVC stimuli, shown in Figure 19, a χ^2 test of association was made between the combinations of false acceptance and refusal errors for CV and CVC stimuli.

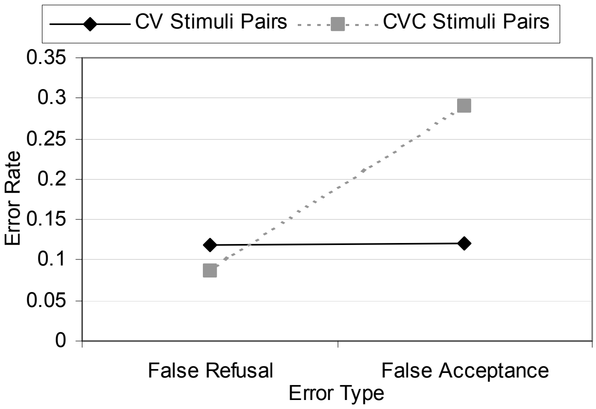


***Figure 19** False Acceptance and False Refusal Vowel Type Differentiation Error rates for CV and CVC stimuli pairs in Experiment 5*

The results of this analysis, shown in Table 10, reveals that false acceptance errors are significantly higher for CVC stimuli pairs than false acceptance errors for CV stimuli, and false refusal errors for both CV and CVC stimuli pairs.

		FA		FR		FR	
		CVC		CV		CVC	
		χ^2	p	χ^2	p	χ^2	p
FA	**CV**	***18.95***	***.0000..***	0.02	.8937	0.44	.5049
FA	**CVC**			***9.82***	***.0017***	***14.17***	***.0002***
FR	**CV**					0.21	.6459

***Table 10** χ^2 Measure of association (df=1, corrected for contingency) between False Acceptance (FA) and False Refusal (FR) error rates for CV and CVC stimuli pairs for Vowel Type Differentiation in Experiment 5*

5.3.5 Conclusions

The primary aim of this experiment was to discover whether subjects were able to perceive the difference between vowels, differentiated only by aperture. Analysis of subject responses from the vowel differentiation task revealed that 81% of all responses were correct, a significant (at $p<0.05$) proportion of correct, over incorrect, responses. These results suggest that subjects are able to distinguish vowel aperture with a significant degree of accuracy. This is similar to that found with the stimuli pairs differentiated by vowel type, that is, vowels that are distinguished by factors other than, or in addition to, aperture, such as roundness, backness, and nasalisation. For this second set of stimuli 82% of all subject responses were correct, a significant (at $p<0.05$) proportion of correct, over incorrect, responses. A comparison of the proportions of correct and incorrect subject responses for vowel aperture differentiation and vowel type differentiation, using a χ^2 test of association, showed that there was no significant difference between the vowel differentiation by type or aperture ($\chi^2_{df(1)} = 0.24$ (corrected for contingency), $p>0.05$). This would suggest that subjects have no more difficulty differentiating vowels by aperture than any other distinguishing feature.

As well as the similarity in overall error rates, subject responses for vowel aperture and vowel type differentiation share another common feature. In both sets of stimuli, an examination of the distribution of false acceptance and refusal errors for CV and CVC stimuli showed that false acceptance errors were significantly higher in CVC stimuli than CV stimuli. As this feature is common across both sets of stimuli pairs it is unlikely that it is related to vowel aperture. It is suggested that the high proportion of false acceptance errors for CVC stimuli is due to the difference in overlap between the two types of stimuli. For stimuli pairs with different vowels CVC pairs have an overlap of two phonemes, the two consonants, but there is only a single overlapping phoneme in CV. It is suggested that CVC pairs have greater similarity than CV pairs, even if they do not share the same vowel, therefore subjects are more likely to perceive them as being identical.

Another feature, found in participant responses for vowel aperture differentiation, was the differences in error rates between vowel types, (/o//ɔ/, /e//ɛ/, and /ø//œ/) for CV or CVC stimuli. It was found that there was a significantly higher occurrence of errors for CVC stimuli with the /o//ɔ/ vowel pair than CV stimuli with the same vowel.

5.4 Summary

The aim of the experiments presented in this section was to verify the validity of the two assumptions underpinning the vowel aperture segmentation hypothesis. This hypothesis states that if the vowel were open then this would encourage a closed syllable, and if the vowel were close this would encourage an open syllable. The first assumption of this hypothesis is that subjects are capable of producing both open and close variants of mid vowels in non-final syllables. The second assumption is that subjects are also capable of perceiving the difference between these vowels.

Although both of these assumptions are theoretically valid, two potential problems were raised in Section 5.1 that could invalidate either, or both, of these assumptions. The first concerns the possibility that, as most of the experimental subjects use the Swiss Romande dialect, that their use of vowel aperture could radically differ from that predicted for ‘standard’ French such that they were limited in the production of either open or close vowels. The second potential problem is that, rather than vowel aperture influencing syllable structure, as suggested by the vowel aperture segmentation hypothesis, syllable structure could influence vowel aperture. In this case open vowels would always be coupled with consonant clusters and singletons that would normally be segmented with a closed first syllable, and vice versa. This would invalidate the vowel aperture hypothesis as an independent syllable segmentation cue, as potential syllabification cues from the vowel aperture would only serve to reinforce the syllabification cues of the intervocalic consonant cluster/singleton.

Because of these potential problems, two experiments were carried out to verify the basic assumptions of the vowel aperture segmentation hypothesis. In the first experiment the production of open and close vowels was tested for words with a single consonant onset. If the first assumption of the hypothesis is valid then subjects should be capable of producing both open and close vowels for these stimuli, whose consonantal structure would usually dictate segmentation with an open first syllable. Analyses of subject vowel production from this experiment would appear to concur with the first assumption. Subjects showed little difficulty in producing words containing a single intervocalic consonant with both open and close vowels.

The second experiment tested the subject's ability to differentiate between CV and CVC stimuli, differing only in the aperture of the vowel. If the second assumption is valid then subjects should be able to differentiate between open and close vowels in either syllabic context. Analyses of subject responses revealed that subjects had little difficulty in differentiating open and close vowels, and that subjects were just as accurate in differentiating open and close vowels as those differing in factors such as nasalisation, roundness, or backness.

These findings suggest that fears surrounding the basic assumptions of the vowel aperture segmentation hypothesis are unfounded. The pool of subjects used in these experiments appears to have little difficulty in either producing or perceiving open and close vowels. There is also no evidence to support the theory that the production of vowel aperture in the non-final syllables is dictated by preferential segmentation dictated by intervocalic consonant cluster/singleton. Therefore, with the basic assumptions of the vowel aperture segmentation hypothesis vindicated, it is possible to proceed onto the segmentation and fragment detection experiments with confidence, in the knowledge that the production and perception of vowel aperture allow its use in syllable segmentation.

6 Vowel Aperture and Syllable Perception

6.1 Introduction

The vowel aperture segmentation hypothesis states that if the vowel of a syllable is open then this will induce syllable segmentation responses that result in the syllable being closed, whereas, if the vowel is close this will induce segmentation responses that result in an open syllable. In the previous section, evidence was found to support the two assumptions of this hypothesis, that listeners are able to both perceive and produce open and close vowels, even if the nature of the intervocalic consonant cluster that follows the vowel dictates a particular syllable aperture (according to the syllable segmentation preference rules suggested by Laporte). The verification of these assumptions ensures that it is likely that there is a suitable phonological environment for the vowel aperture segmentation hypothesis. The aims of this section are to investigate whether the behaviour predicted by this hypothesis is evident in the segmentation behaviour of naïve listeners.

Two experimental paradigms are used to investigate the validity of the vowel aperture segmentation hypothesis. The first uses a metalinguistic task, to investigate the broader implications of the syllabification cues arising from vowel aperture and the phonotactics of the intervocalic consonant cluster or singleton. In the second, an on-line task is used to investigate whether the locus of the proposed hypothesis is more likely to be based on the pre-lexical or post-lexical phonological level.

The first experiment uses the metalinguistic syllable repetition task employed in Experiment 1. However, in this instance, as well as controlling the intervocalic consonant cluster/singleton of the stimuli, the aperture of the first vowel is also controlled. This will allow the examination of the effect of vowel aperture upon syllable segmentation by comparing the proportion of open and closed segmentation responses for each vowel aperture.

In the second experiment a subset of the /VCCV/ stimuli from the previous experiment are used as carrier items in a fragment detection task. The aim of this experiment is to reproduce the syllable effect found in the fragment detection task of Mehler et al (1981). However, in a departure from the original experimental design, instead of referencing the syllable structure of the carrier from the nature of the intervocalic consonant cluster/singleton, it is referenced using the vowel aperture segmentation hypothesis. Therefore, in this case, the syllable effect would be represented as an interaction between VC and C targets and vowel aperture, where VC targets should be detected faster than V targets in carriers with open vowels, and V targets faster than VC targets in carriers with close vowels.

6.2 Experiment 6 – Examination of the effect of Vowel Aperture in Syllable Repetition

The primary aim of this experiment is to investigate the vowel aperture segmentation hypothesis. As has been previously discussed, it has been suggested that if the aperture of a consonant cluster/singletons preceding vowel is close, this will encourage segmentation responses resulting in an open first syllable, whilst an open preceding vowel will encourage a closed first syllable. To investigate this hypothesis a syllable repetition task was implemented, using a subset of the consonant clusters and singletons used in Experiment 1. These clusters were chosen to reflect a cross section of syllabification behaviour found in Experiment 1 (three OBLI clusters, the /ks/ cluster, and the six remaining clusters split between non-OBLI clusters with one or two legal onsets). When generating stimuli for Experiment 1 only the consonant cluster/singleton was controlled, with the bracketing vowels selected at random, in Experiment 6 the first vowel of the stimuli was also controlled.

When creating stimuli, each of the open and close variants of the three mid vowel types will be combined with each of the consonant clusters/singletons. This will allow the examination of the effect of vowel aperture upon syllable segmentation by comparing the proportion of open and closed segmentation responses for each vowel aperture.

6.2.1 Stimuli

6.2.1.1 Test items

Non-word stimuli were generated using a subset of consonant clusters/singletons used in Experiment 1. Consonant cluster/singletons used in this experiment were composed of two cluster/singletons for each of the N, F, L, P, FL, PF, FL, FP, and LN robust feature classes (in order to maintain a similar balance of stimuli as found in Experiment 1). Each of these consonant clusters/singletons were appended to each mid-vowel with open/close counterparts (/o//ɔ/, /e//ɛ/, and /ø//œ/). Final vowels were randomly selected from [/u/, /i/, /a/, /ǎ/ /y/] with open/close agreement between the first and second vowel. However, in a few cases these constraints could not be satisfied, as all available choices for the final vowel resulted in a lexical stimulus. In these cases, [ɔ] was used as an alternative for the final vowel. This organisation resulted in similar stimuli to those used in Experiment 1, with 6 tokens resulting from each of the 18 consonant clusters/singletons. A full list of stimuli appears in Appendix F.

6.2.1.2 Distractor and Training items

To maintain a similar distribution of stimuli types to those found in Experiment 1, a number of VCCCV stimuli were included as distractors using the /rbl/ and /kst/ consonant clusters. These items were constructed using the same constraints as the test stimuli, resulting in 6 tokens for each of the triple consonant clusters.

A total of 10 training items were also generated, in similar form to that of the test items and using the same distribution of VCV, VCCV, and VCCCV stimuli used in the test and distractor sets, but having different consonant cluster/singletons.

All stimuli were produced by a female adult monolingual, naïve (that is, they had no knowledge of the aims or procedure behind the experiment), native French speaker (Parisien accent) from a randomised list of non-words represented phonetically using the IPA alphabet, with a pause of three seconds inserted between the production of each stimulus. These non-words were recorded using a Sennheiser ME80 microphone and Rane MS1 microphone pre-amplifier onto Sony ZA5ES DAT recorder. Afterwards the stimuli were low-pass filtered at 11kHz and then digitised with a fidelity of 22kHz with 16 bits per sample.

6.2.2 Procedure

As Experiment 1.

6.2.3 Participants

All 36 participants (34 female, 2 male) were students of the Université de Genève, with an average age of 22;1 (ranging from 18 to 39), and were native speakers of French with no known hearing defects. They received course credits for their participation.

6.2.4 Results

6.2.4.1 Error rates

Errors consisted of missing responses and repetition of the total stimulus (both syllables). Other sources of error were the mispronunciation of syllable repetitions, that is, the use of phonemes not found in the original stimuli (excepting vowel aperture errors, see below). Error rates averaged at just fewer than 2% of the possible responses. There was no significant effect of any experimental factor on error rates, with the highest subject error rate at 6.4%.

Mispronunciation 'errors' involving a disparity between the vowel aperture of the first vowel of the stimulus, and the subject response in experimental condition 1 (repetition of the first syllable) were not included in these errors. For example, in experimental condition 1, the subject is presented with the stimulus /ɛgla/, and responds with the utterance /eg/ instead of /ɛg/ the disparity in the first vowel aperture between stimulus and response is not counted as an error. This is because there may be a bias for the production of open vowels with closed syllables and vice versa, inline with the usual distribution of open and close vowels and syllable aperture in normal spoken French (as summarised in Section 4.3). Therefore, it is possible that these disparities may be due to the influences of speech production, and not perception. That is, they could reflect the syllable aperture of the response (open vowels produced with closed syllable responses and vice versa), and not the perceived aperture of the first vowel. This disparity between the first vowel aperture and subject responses occurred in 27% of all subject responses in experimental condition 1.

6.2.4.2 Preferential Segmentation of Consonant Cluster/Singletons

Setting aside the possible influences of the first vowel of the stimuli, segmentation responses were analysed across each cluster/segment to test for the presence of a significant preferential segmentation. Responses were analysed separately for each of the experimental conditions (first and second syllable repetition). For each cluster/singleton type a χ^2 was calculated on the frequencies of possible responses to determine whether there was a preferential category of response.

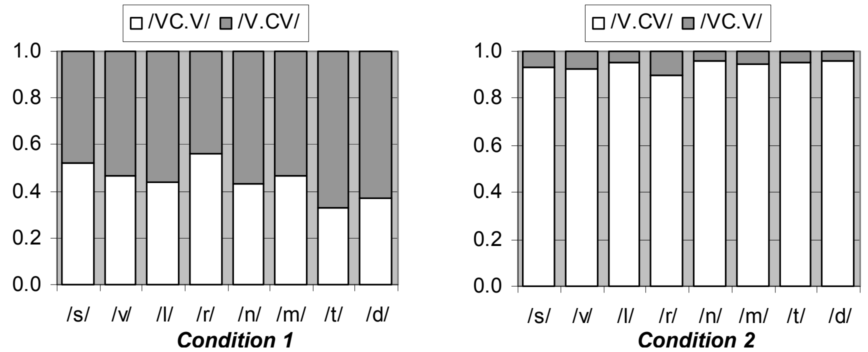


***Figure 20** Distribution of subject syllabification responses in Experiment 6 for experimental conditions 1 and 2 for all single consonant stimuli*

An examination of the distribution of segmentation responses for single intervocalic consonant stimuli, as seen in Figure 20, shows a similar distribution of responses to those found in Experiment 1. With the high proportion of increased onset ambisyllabic responses there was no significant preferential response ($p>0.05$) for responses from experimental condition 1 for the /s/, /v/, /l/, /r/, and /m/ singletons. Only for /n/, /t/, and /d/ singletons was the χ^2 higher than the critical level (at p= 0.05, 3.841 (df=1)) with a /V.CV/ preferential segmentation. Responses from experimental condition 2 showed a significant preferential segmentation of /V.CV/ for all singletons (at p= 0.05, 3.841 (df=1)).

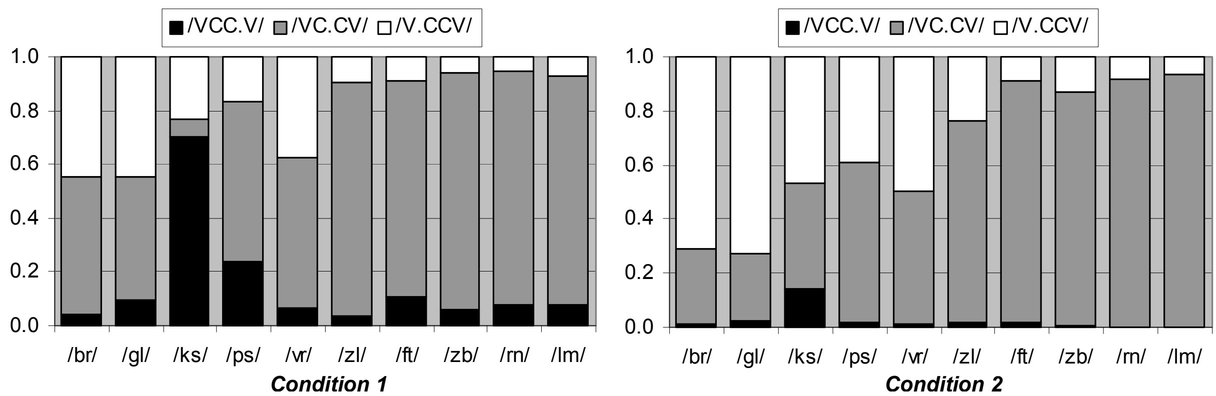


***Figure 21** Distribution of subject syllabification responses in Experiment 6 for experimental conditions 1 and 2 for all double consonant stimuli*

An examination of the distribution of subject responses for double consonant stimuli, shown in Figure 21, reveals a number of differences in the preferential segmentation of double intervocalic consonant cluster stimuli between experiments 1 and 6. For all consonant clusters the χ^2 was higher than the critical level (p=0.05, 5.991 (df=2)) when calculated on all three of the possible subject responses (/V.CCV/, /VC.CV/, and /VCC.V/). However, in the case of the responses from Experiment 6, this analysis does not give a true depiction of significant preferential segmentation, due, in the majority of cases, to the relative paucity of /VCC.V/ responses which could be considered as an artefact. For example, in the case of subject responses for the /vr/ cluster in experimental condition 2, there were 103 responses for both /.vr/ and /v.r/ segmentations, but only 2 for /vr./. In this case, although the χ^2 measure based upon all three possible responses is above the critical level it is clear that there is no preferential segmentation, with the majority of responses split evenly between single and double syllable onset segmentation. The cluster that provides the exception to this problem is that of /ks/, with the preferential segmentation (70% of subject responses) being /ks./ in condition 1. In this case, as previously explained in Experiments 1 and 3, the unusual distribution of subject responses for this cluster is considered to be due to orthographic influences. For the remaining consonant clusters, and for /ks/ in experimental condition 2, analyses were repeated taking only /V.CCV/ and /VC.CV/ segmentations into account. Examining responses from experimental condition 1, all clusters apart from /br/ and /gl/ resulted in a significant preferential segmentation with the χ^2 higher than the critical level (p=0.05, 3.841 (df=1)). All significant preferential segmentations follow those found in Experiment 1, with all clusters segmented as /VC.CV/. In experimental condition 2 all clusters apart from /ks/ and /vr/ had significant preferential segmentation (p=0.05, 3.841 (df=1)), with segmentation following those found in Experiment 1, with OBLI clusters segmented as /V.CCV/ and those remaining as /VC.CV/.

6.2.4.3 <u>Analysis of the Influence of the Vowel in the Segmentation of Single Intervocalic Consonant Stimuli</u>

To analyse the possible influence of the aperture of a single intervocalic consonant's preceding vowel upon its segmentation a loglinear analysis was conducted for each experimental condition. Subject responses were represented by the frequency of /V.CV/ responses over two within subject factors, consonant type (8 consonants) and vowel aperture (two modalities, open or close). Missing data was replaced with the average value of the subject responses for each particular stimulus.

In experimental condition 1, the analysis revealed that the best model had a generating class with main effects of vowel aperture and consonant (Likelihood Ratio χ^2 = 3.03, df = 7, p=0.882). The main effect of the consonant follows the behaviour found with single intervocalic consonant stimuli in Experiment 1, covered in Section 3.3.5.1, and also shown in Figure 20 for this experiment, with sonorants more closely linked to the preceding vowel than obstruents. In Experiment 6, 64% of subject responses for stimuli containing plosive consonants, resulted in a /V.CV/ segmentation, with 49% for liquid consonants. The main effect of aperture described a higher proportion of /V.CV/ segmentation responses for stimuli with close vowels than for open vowels. For stimuli with open preceding vowels, 48.67% of responses resulted in an open first syllable, whilst for stimuli where the preceding vowel was close there were 61.20% of responses for an open first syllable.

In experimental condition 2, the analysis revealed that the best model had a generating class with no factors (Likelihood Ratio χ^2 =1.62, df = 15, p=1.0). This result confirms that found in Experiment 1, showing that, whilst the sonority of the intervocalic consonant had an effect upon syllable boundary placement in syllable offset detection, there was no effect in syllable onset detection.

6.2.4.4 Analysis of the Influence of the Vowel in the Segmentation of Double Intervocalic Consonant Stimuli

Similar to that performed in the previous section, loglinear analyses were used to examine subject responses for double intervocalic consonant stimuli. Subject responses were represented by the frequency of /V.CCV/ responses over two within subject factors, consonant cluster type (10 clusters) and vowel aperture (two modalities, open or close). As the primary aim of this experiment was the investigation of a possible link between open vowels and closed syllables, and vice versa, it is not necessary to make the distinction between the two types of segmentation which result in a closed first syllable. Also, apart from the cluster /ks/, subject responses for the /VCC.V/ segmentation generally constitute only a very small minority of all subject responses. Missing data was replaced with the average value of the subject responses for each particular stimulus.

In experimental condition 1, the analysis revealed that the best model had a generating class with main effects of aperture and consonant cluster (Likelihood Ratio χ^2 = 7.81, df = 9, p=0.553). The main effect of the consonant describes the preferential segmentation responses shown in section 6.2.4.2 and follows the behaviour found with double intervocalic consonant stimuli in Experiment 1. The direction of the effect of vowel aperture was the same as that found with single intervocalic consonants, 17% of responses for stimuli with open vowels resulted in an open first syllable, with 23% of responses for close vowels.

In experimental condition 2, the analysis revealed that the best model had a generating class with only the main effect of consonant cluster (Likelihood Ratio χ^2 = 6.99, df = 10, p=0.726). The main effect of the consonant describes the preferential segmentation responses shown in section 6.2.4.2 and follows the behaviour found with double intervocalic consonant stimuli in Experiment 1.

6.2.4.5 Analyses of Vowel Aperture Errors

As mentioned in Section 6.2.4.1, errors in the production of vowel aperture in subject responses from experimental condition 1 were not treated as experimental errors, and so, were not excluded from analysis. These responses included due to the high incidence of this type of error (27% of all responses from experimental condition 1), and also because of uncertainty surrounding the nature of this apparent error.

The origin of vowel aperture errors could be based on either, perception, where the subjects perceive the vowel aperture incorrectly, or production, where the vowel aperture is correctly identified, but not correctly produced in the subject response. If the errors were based on production, then this would make no difference to the measured effect of vowel aperture upon syllabic segmentation. As the main interests of this study lie with syllable perception, rather than production, these errors are of limited interest. However, if the errors were due to problems in perception then they would increase the possibility of type II errors in the analyses of the effect of vowel aperture since the analysis of results would not be based upon the correct, perceived, vowel aperture of all stimuli. A certain degree of vowel aperture perception errors are expected, as was found in Experiment 5, where 18% vowel aperture differentiation responses were in error. Fortunately, although the existence of such errors may have distorted the influence of vowel aperture upon syllable segmentation, even with such errors the main effect of vowel aperture was found across all stimuli in experimental condition 1.

Two hypotheses are presented for the explanation of these errors, the production and perception hypotheses, each based upon the relationship between vowel aperture errors and syllabification responses. These hypotheses, however, are not considered mutually exclusive, as it is possible that vowel aperture errors could result from a mix of both production and perception errors.

The production hypothesis, suggests that vowel aperture is perceived correctly, but the production of the vowel aperture is related to the syllabification response given by the subject. Therefore, it is suggested that subjects are more likely to use an open vowel with a closed syllable response, and a close vowel with an open syllable response, even if the vowel aperture of the stimuli is not in agreement with the subject response. For example, presented with the stimuli /ebri/ the subject is more likely to respond with /ɛb/ rather than /eb/ for a /VC.CV/ segmentation response since the subject is matching the vowel aperture with the response, not the stimulus. If this hypothesis is correct, then subjects should produce fewer vowel errors when there is agreement between the stimulus vowel aperture and the response (open vowel stimulus with /VC.CV/ or /VCC.V/ responses, or close vowel stimulus with /V.CCV/ response), than where there is disagreement (open vowel stimulus with /V.CCV/ response, or close vowel stimulus with /VC.CV/ or /VCC.V/ responses).

To test this hypothesis, stimuli were split into two categories, namely, stimuli using open vowels, and those using close vowels. For each of the stimuli in these categories the percentage of open syllable responses (/V.CCV/ or /V.CV/) was compared with the percentage of vowel production errors, shown in Figure 22. If the hypothesis is valid there should be a positive correlation between open syllable responses and open syllable vowel aperture errors in the first set, and a negative correlation between open syllable responses and closed syllable vowel aperture errors in the second set. Results of these comparisons appear to support the production hypothesis, with a significant positive correlation for the comparison of open vowel stimuli ($r = .617$, $t(52)= 5.648$, $p<0.05$), and a significant negative correlation with close vowel stimuli ($r = -.727$, $t(52)=-7.635$, $p<0.05$).

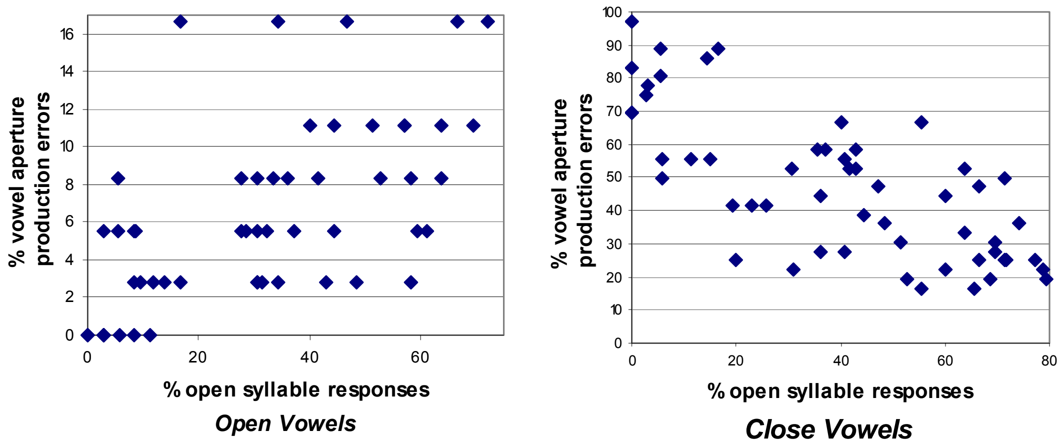


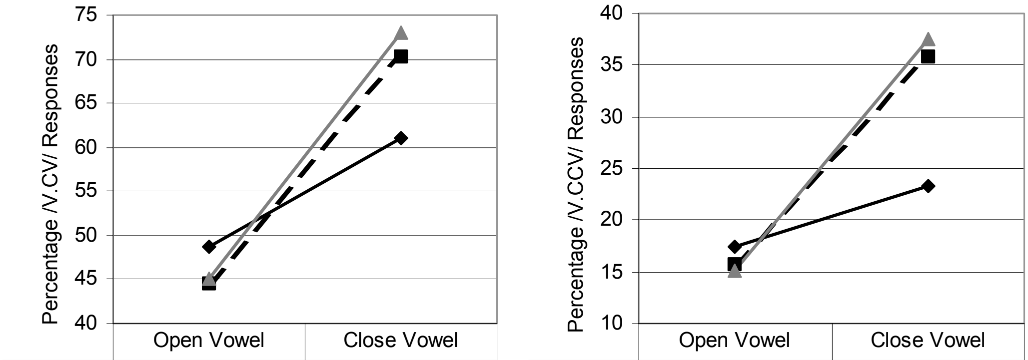
Figure 22 *Comparison of the percentage of vowel aperture production errors and open syllable responses for open and close vowels*

These results would suggest that the choice of vowel aperture used in subject responses is not always in agreement with that used in the stimulus, but is also influenced by the syllabic nature of the response. It would appear that there is a tendency amongst subjects to produce responses that are in agreement with the normal phonological rules of French. This would encourage the use of close vowels when uttered in isolation and open vowels when they are appended with a coda.

The second hypothesis, the perception hypothesis, suggests that the errors in vowel aperture production could be due to errors in the perception of the vowel aperture, with subjects repeating the vowel aperture that they perceive, whether it is the correct aperture for the stimulus or not. If this hypothesis is valid then the effect of vowel aperture, found in sections 6.2.4.3 and 6.2.4.4, will be diluted, since it is measured on the aperture of the stimulus, not on that produced by the subjects. For subject responses with the correct vowel aperture perception the direction of the aperture effect will show an increase in open syllable responses for close vowels, and vice versa. Conversely, because responses with incorrect vowel aperture perception are referenced from the vowel aperture of the stimulus, there should still be the same effect of vowel aperture, but in the opposite direction as the perception of the vowel aperture is the opposite to that used in the stimulus.

To test this hypothesis subject responses were separated into two categories, those with correct vowel aperture production, and those with incorrect vowel aperture production. For each category a test of correlation was made between the mean percentage of open syllable responses for each stimulus and the vowel aperture of the stimulus (represented in a binary form, 0 for open vowels, 1 for close vowels). If the perception hypothesis is correct then there should be a positive correlation between the percentage of open syllable responses and vowel aperture (as close vowels are represented by 1) for the category of correct vowel production responses, and a negative correlation for the category of incorrect vowel production responses. The results of these analyses appear to support the perception hypothesis, with a significant positive correlation found between the percentage of open syllable responses and vowel aperture for the category of correct vowel aperture production responses in both single ($r = .702$, $t(46)=6.704$, $p<0.05$) and double consonant stimuli ($r = .358$, $t(57)=2.903$, $p<0.05$). For the category of incorrect vowel aperture production responses there was a significant negative correlation, for both single ($r = .710$, $t(46)=6.846$, $p<0.05$) and double ($r = -0.591$, $t(49)=5.138$, $p<0.05$) consonant stimuli.

These results suggest that vowel aperture errors were caused, not only by errors in production, but also by the perception of vowel aperture. This finding raises questions about the accuracy of the degree of influence that vowel aperture has over syllable segmentation, since these analyses included subject responses where the perception of the vowel aperture was not in agreement with that of the stimuli. This could lead to a distortion of the effects and interactions of these analyses involving the factor of vowel aperture. Whilst there was a main effect of vowel aperture in experimental condition 1, the difference between the segmentation behaviour of stimuli with open and close vowels is reduced by the inclusion of these erroneous responses. This can be seen in Figure 23, where the difference in the percentage of open syllable responses due to vowel aperture is greater for subject responses where there is vowel aperture agreement between stimuli and responses, than for all responses, with vowel aperture referenced from the stimuli. Also, if the mean percentage of all open syllable responses is recalculated, referencing the vowel aperture, not by that used in the stimulus, but produced by the subject, the results are similar to the subset of responses where there was aperture agreement between stimuli and subject response. These responses may provide a more accurate representation of the true extent of the influence of the vowel aperture over syllabification; with a 25% swing (between stimuli with open and close vowels) in open syllable segmentation decisions for single consonant stimuli, and 20% for double consonant stimuli.



	Open Vowel	Close Vowel
1	48.64	61.12
2	44.57	70.24
3	44.99	72.96

Single Consonant

	Open Vowel	Close Vowel
1	17.38	23.26
2	15.73	35.82
3	15.09	37.5

Double Consonant

Figure 23 *Percentage of open syllable responses in condition 1 of Experiment 6 for single and double consonant stimuli (1 – All responses, vowel aperture referenced from stimuli, 2 – Responses with agreement in vowel aperture between stimuli and response, 3 – All responses, vowel aperture referenced from subject response)*

6.2.5 The Effect of Vowel Aperture and Vowel Aperture Production Errors

Repetition errors involving vowel aperture, referred to as vowel aperture errors, describe subject responses where there was disparity between the vowel aperture produced in the response and that of the stimuli. These errors were accorded special status, set apart from the other types of repetition errors, for two reasons. Firstly, a high proportion of subject responses containing these errors, far higher than other repetition errors (in experimental condition 1, 27% of responses contained vowel aperture errors, 2% for other experimental errors). Second, the nature of these errors is ambiguous, and can be attributed to errors in the perception of the vowel aperture, or, that subjects are changing the aperture of the vowel in their response to coincide with aperture of their syllabic response (open or closed syllable). By including subject responses with vowel aperture errors in the analyses of syllable segmentation a comparative analysis of segmentation and vowel aperture errors can be made, which can reveal the impact of these errors upon syllable segmentation, and also the nature of these errors.

Two hypotheses were presented to explain these findings, and also to examine the nature of vowel aperture errors. The production hypothesis stated that subject would tend to produce the vowel aperture that was in agreement with the aperture of the their syllabic response, using an open vowel with a closed syllable, and vice versa. The perception hypothesis suggested that errors in vowel aperture were due to errors in vowel perception, and that the vowels that subjects used in their responses were those that they had perceived. Comparison of the distributions of vowel aperture errors and subject segmentation responses revealed that there was evidence to support both of these hypotheses, suggesting that the mismatches between the vowel aperture used in the subject response and that of the stimuli were due to factors of both perception and production. Whilst the vowel aperture errors due to production have no bearing upon the segmentation results, errors in perception will distort the effects of vowel aperture found in the original analyses of syllable segmentation.

Although there was a significant effect of vowel aperture across all stimuli types in experimental condition 1 the difference between the mean number of open syllable responses for open and close vowels was minor. Averaging across all stimuli and experimental conditions there was only a 9% increase in open syllable segmentation responses for close rather than open vowels. This figure is based upon analyses of subject responses where the vowel aperture is referenced from the stimuli, which may be considered to be subject to distortion due to vowel aperture perception errors. However, if the vowel aperture is referenced from the response produced by the subject, reducing the effect of vowel aperture perception errors, the effect of vowel aperture increases substantially, with a 30% increase in open syllable segmentation responses for close rather than open vowels.

6.2.6 Conclusions

The primary aim of this experiment was to investigate whether the aperture of a vowel could have any influence over the syllabification of the intervocalic consonant cluster or singleton that proceeded it. It had been hypothesised that there was a link between the aperture of the vowel at the syllable’s nucleus, and the aperture of the syllable, with open vowels encouraging closed syllables and vice versa.

Analyses of open syllable segmentation responses from Experiment 6 appear to support this hypothesis, with a significant effect of vowel aperture found for the subject responses from both single and double consonant stimuli for the first experimental condition (first syllable repetition).

The direction of this effect is that which had been hypothesised, with a greater number of open syllable responses for stimuli with close vowels than those with close vowels.

However, as the effect of vowel aperture was only found in the first experimental condition, linked to syllable offset detection, the evidence in support for the vowel aperture segmentation hypothesis is inclusive. If, as suggested by Content et al (in press), syllable segmentation were driven purely by syllable onset detection, then vowel aperture would have no effect on the segmentation of syllables at the pre-lexical processing level. Therefore, to address this ambiguity it is necessary to find additional evidence to support the vowel aperture segmentation hypothesis using an 'on-line' experimental paradigm, designed to reveal the workings of syllabic segmentation at a pre- rather than post-lexical level of processing.

6.3 Experiment 7 – Fragment Detection and the Syllable Effect

As has been discussed in Section 1.1, one of the most important arguments for the use of the syllable as a segmentation unit of speech in French comes from evidence favouring the hypothesis that it is the syllable, rather than the phoneme, that is used at the pre-lexical level of speech processing.

Possibly the most influential source of evidence surrounding the use of the syllable in pre-lexical processing come from studies using the fragment detection experimental paradigm. In the original study, by Mehler et al (1981), French subjects detected CV or CVC targets in spoken carrier words that varied in syllabic structure. Carrier pairs each shared the same initial three phonemes (CVC), and were selected such that these phonemes made up the first syllable of the first member of the pair, and that the first two phonemes (CV) formed the first syllable of the second member of the pair (e.g. /pa.lace/ and /pal.mier/). Detection latencies showed that there was a cross-over interaction between target type and carrier syllable structure, such that responses were faster when the target corresponded to the initial syllable of the carrier. The rapidity of detection latencies in this task, at around 360ms, is thought to eliminate the possibility of post-lexical access responses, and therefore it was suggested that the locus of this 'syllable effect' was pre- rather than post-lexical.

The aim of Experiment 7 is to reproduce the syllable effect found in the fragment detection task of Mehler et al (1981). In the original fragment detection task the syllable structure of the carrier items was referenced from the phonotactics of the intervocalic consonant and singleton used in the carrier. In Experiment 7, the syllable structure of the carriers is referenced, not by the nature of the intervocalic consonant cluster/singleton, but by the aperture of the first vowel. In Experiment 6, a metalinguistic syllable repetition task revealed evidence to support the hypothesis that open vowels induce closed syllables, and close vowels induce open syllables. The purpose of Experiment 7 is to verify these findings using an on-line experimental paradigm, the fragment detection task.

In this experiment, carrier items consist of the stimuli used in Experiment 6, containing the /lm/, /ps/, /gl/, and /vr/ intervocalic consonant clusters. Using the syllable segmentation proposed by the vowel aperture segmentation hypothesis, these carriers are arranged in pairs according to the aperture of their first vowel. Target items will consist of V and VC fragments, such that V target items will make up the first syllable of carriers with close first vowels, and VC target items will make up the first syllable of carriers with open first vowels.

Therefore, if there is a syllable effect as a result of the segmentation predicted by the vowel aperture segmentation hypothesis, VC targets should be detected faster than V targets in carriers with open vowels, and V targets faster than VC targets in carriers with close vowels. An alternative hypothesis is that of phonemic matching, with the target detected as soon as all of its phonemes are identified in the signal. In this case V targets will be detected more rapidly than VC targets for carriers with both open and close vowels.

6.3.1 Stimuli

6.3.1.1 <u>Test Stimuli</u>

Carrier non-words consisted of a subset of stimuli from Experiment 6 using the intervocalic consonant clusters, /lm/, /gl/, /vr/, and /ps/. This selection resulted in six carrier items for each cluster, one for each of the /o/, /ɔ/, /e/, /ɛ/, /ø/, and /œ/ first vowels. For each of the carriers a V and VC target were used, consisting, respectively, of the first vowel, and first vowel and consonant, of the carrier item. For a full transcription of the carrier items, see Appendix F.

6.3.1.2 <u>Foils</u>

For each carrier, two foil ‘targets’ were generated, a V and VC foil. The V foil consisted of an isolated vowel, chosen from any of the mid vowels, excepting that found used as the first vowel of the carrier (e.g. for the carrier /ɛgla/ the foil /o/ was used). The VC foil consisted of a vowel, matching the first vowel of the carrier, followed by a consonant, different to that of the second consonant of the carrier (e.g. for the carrier /ɛpsa/ the foil /ɛm/ was used). The purpose of the VC foil is to ensure that subjects could not respond on the basis of a partial match on the first segment of the target.

6.3.1.3 <u>Training Items</u>

A total of 10 training carrier items were selected from the double consonant cluster stimuli of Experiment 6 not used in the test carrier items. As in the test stimuli, for each carrier V and VC targets and foils were also created. This resulted in a total of 20 target carrier, and 20 foil carrier training items.

A different monolingual French speaker was used to produce the target and foil items to that used in the production of the carrier items (in Experiment 6) to avoid the possibility of simple target carrier acoustic matching. Both speakers were trained phoneticians with a Parisien accent, of the same sex (female), and had no knowledge of the aims or procedure behind the experiment. Stimuli were recorded using a Sennheiser ME80 microphone and Rane MS1 microphone pre-amplifier onto Sony ZA5ES DAT recorder. Afterwards the stimuli were low-pass filtered at 11kHz and then digitised with a fidelity of 22kHz with 16 bits per sample.

6.3.2 Procedure

The experiment was performed automatically using the Psyscope experimental software running on a Macintosh computer with accompanying button box. Audio stimuli were reproduced using this software through an audiomedia sound card and presented to subjects using Beyer Dynamic DT100 headphones. During the experiment subjects were seated in an amplisilence acoustic booth.

Subjects were asked to press a button as rapidly as possible if the target stimulus was present at the beginning of the carrier. Target stimuli were presented after a short warning tone, followed by the carrier stimulus 300ms afterwards. Each subject received all of the carriers four times, once for each of the V target and VC targets, and once each of the V and VC foils. Target and carrier stimuli pairs were presented in random order without pause, with one pair presented every two seconds. At the start of the experiment a training block was presented, consisting of 20 target and carrier stimuli pairs and 20 foil and carrier pairs presented in random order.

6.3.3 Participants

All 19 participants (18 female, 1 male) were students of the Université de Genève, with an average age of 23;5 (ranging from 19 to 39), and were native speakers of French with no known hearing defects. They received course credits for their participation.

6.3.4 Results

6.3.4.1 Error Rates

Errors for test items consisted of missed responses, which accounted for 2.2% of all expected responses. In addition, reaction times of longer than 1000ms and shorter than 100ms were excluded from further analysis. The excluded data made up less than 2% of all subject responses. Subjects erroneously detected 7.3% of the foil targets in carrier items. All errors, both in test stimuli and foil items, were randomly distributed amongst carriers and targets/foils.

6.3.4.2 Analyses of Reaction Time Differences for V and VC Targets

To analyse the variation in reaction times for V and CV targets between target carrier pairs with open and close vowels a repeated measure ANOVA was conducted with four within subject factors, cluster type (/gl/, /lm, /vr/, and /ps/), vowel type (three types, /ø//œ/, /o//ɔ/, /e//ɛ/), vowel aperture (two modalities, open or close), and target type (V or VC). Missing data was replaced with the average cross subject reaction time for each particular target carrier pair.

The analysis revealed that there were no significant main effects for any of the experimental factors, but that there was a significant interaction between target type and vowel aperture ($F(1,18)_{0.05} = 14.77$). This interaction, shown in Figure 24, reveals that subjects were faster in detecting V targets in target carrier pairs with close vowels than VC targets, with the reverse for target carrier pairs with open vowels.

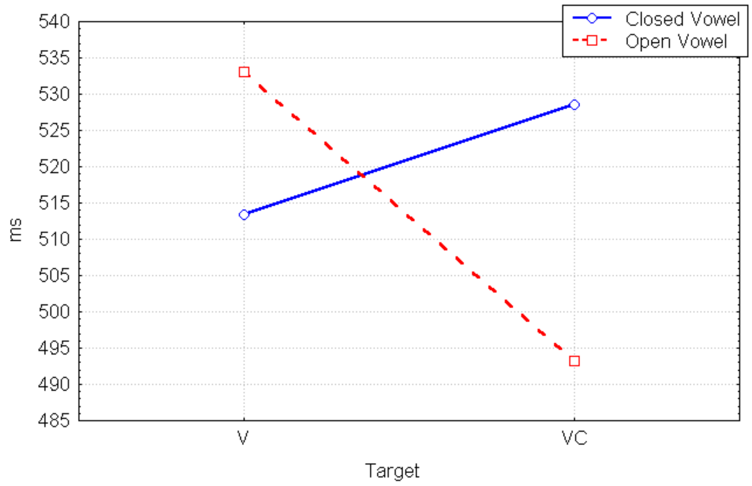


Figure 24 *V and VC target reaction times for target carriers pairs using open and close vowels in Experiment 7*

Although this interaction would appear to support the vowel aperture segmentation hypothesis, an inspection of the results also reveals that the difference in reaction times between V and VC targets with close vowels (15ms) is smaller than the difference between VC and V targets with open vowels (50ms). Accordingly, separate ANOVA analyses of reaction times were conducted for each vowel aperture. These analyses revealed that, whilst there was a significant main effect of target for target-carrier pairs with open vowels ($F(1,18)_{0.05} = 12.87$), the main effect of target for target-carrier pairs with close vowels did not approach significance ($F(1,18) = 2.45$, $p = 0.10$).

6.3.4.3 <u>Reaction Times for Fast and Slow Subjects</u>

In the original fragment detection task by Mehler et al (1981) it had been suggested that the syllable effect was pre- rather than post-lexical because the latency of the reaction times, averaging around 360ms, was thought to be too rapid for lexical access. However, in the current experiment average reaction times are considerably slower, at 517ms (averaged over all carrier and target pairs), which, coupled with an average carrier length of 530ms, may not eliminate the possibility of a post-lexical access response. A possible explanation for the length of these reaction times might lie in the use of VC foils in this experiment, which were not used in the original study. In study of the syllable effect in French, using non-word stimuli, Meunier et al (1997) found that the use of CVC foils increased average reaction times by approximately 100ms when compared with a similar experiment where only CV foils were used.

To analyse the possible differences between slow and fast responses, subjects were split into two groups. The first group consisted of the nine subjects with the fastest average reaction times (average reaction time of 443ms), the second consisting of the remaining ten subjects with slower reaction times (average reaction time of 591ms). The ANOVA analyses used in the previous section were repeated for each of the subject groups.

Fast Subjects

ANOVA analysis of the reaction times of the nine fastest subjects revealed a similar pattern of responses to those found with all subjects. As before, a significant interaction was found between vowel aperture and target type ($F(1,8)_{0.05}$ =10.19), but no main effects for any of the experimental factors. The direction of this interaction was the same as that found in the previous analysis. Separate ANOVA analyses for open and close target carriers pairs revealed that there was a significant main effect of target for target-carrier pairs with open vowels ($F(1,8)_{0.05}$ =8.20), and that the main effect of target for target-carrier pairs with close vowels was not significant ($F(1,8) = 2.33$, $p = 0.165$).

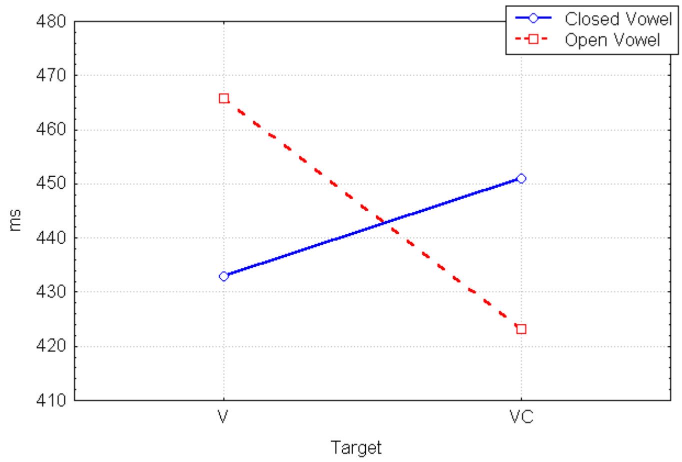


Figure 25 *V and VC target reaction times for target carriers pairs using open and close vowels in the fast subject group of Experiment 7*

Slow Subjects

ANOVA analyses for the reaction times of the ten slowest subjects revealed significant differences in the distribution of detection latencies amongst the experimental factors between the fast and slow group of subjects. Firstly, an examination of the main effects of the experimental factors showed a significant main effect of vowel aperture, ($F(1,9)_{0.05}$ =56.64), with detection latencies smaller for target stimuli pairs with open vowels (580ms) than close vowels (610ms). In addition, the interaction between vowel aperture and target type shown in the previous analyses was found to be insignificant ($p>0.05$) in the group of slow subjects. As can be seen in Figure 26, although there is a similar interaction to that found in fast subjects, this is significantly weaker for the slower subjects, with the loss of significance possibly due to a high deviation amongst reaction times. Therefore, the evidence suggest that, for slow subjects, with average reaction times longer than the mean length of carrier items, the possible use of metalinguistic knowledge *reduces* the syllable effect caused by vowel aperture.

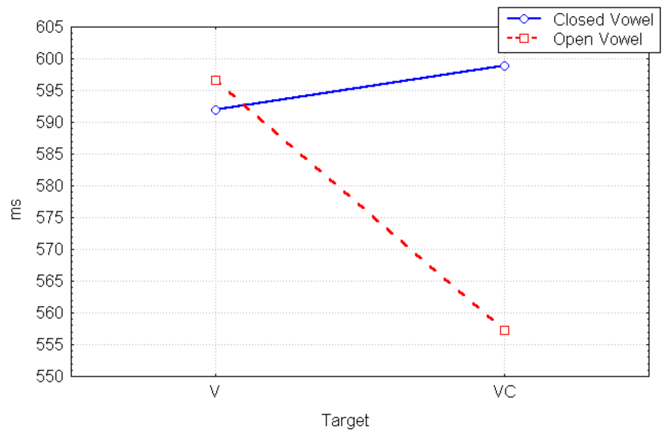


***Figure 26** V and VC target reaction times for target carriers pairs using open and close vowels in the slow subject group of Experiment 7*

6.3.5 Discussion

As in the original fragment detection task from Mehler et al (1981), detection latencies showed a crossover interaction between target type and interaction, with faster responses when the target corresponded to the initial syllable of the carrier. However, in a departure from the original experiment, in this case the syllable structure was not referenced from the phonological structure of the intervocalic consonant cluster/singleton, but from the aperture of its preceding vowel. With V targets detected faster than VC targets with close vowels, and VC targets faster than V targets with open vowels, these results would appear to support the vowel aperture segmentation hypothesis. That is, that close vowels induce open syllables, and open vowels induce closed syllables.

A comparison of the differences in the distribution of reaction times for fast and slow subjects may also suggest the locus of this effect, pre- or post-lexical. For fast subjects, with an average reaction time of around 430ms, the interaction between vowel aperture and target type is significant. However, for slow subjects, with an average reaction time of 591ms, longer than the average length of the carrier items, there is no significant interaction between vowel aperture and target type. These results suggest that the locus of the syllable effect caused by vowel aperture is more likely to be at the pre-lexical rather than lexical or post-lexical level. This is because the syllable effect reduces as average reaction times increase to a level that would allow a post-lexical access response.

In final conclusion, the results of the fragment detection task appear to confirm the evidence found in Experiment 6 supporting the vowel aperture segmentation hypothesis.

6.4 General Discussion

The vowel aperture segmentation hypothesis states that if the vowel of a syllable is open, then this will induce syllable segmentation responses that result in the syllable being closed, and if the vowel is close, this will induce segmentation responses that result in an open syllable. The aim of this section was to investigate whether the behaviour predicted by this hypothesis is evident in the segmentation behaviour of naïve listeners.

Two experiments were used to investigate the claims made by the vowel aperture segmentation hypothesis, a syllable repetition experiment, and a fragment detection experiment. Evidence from the subject responses in both experiments appears to support the hypothesis that vowel aperture is used as a segmentation cue by listeners.

In Experiment 6 a syllable repetition task was performed, similar to that used in Experiment 1. However, whilst in Experiment 1 only the intervocalic consonant cluster/singleton of the stimuli was controlled, in Experiment 6 the first vowel of the stimuli was also controlled, with six stimuli tokens generated for each consonant cluster/singleton under analyses, one for each of the mid vowels. Analyses of the segmentation decisions of the experimental subjects revealed a significant effect of vowel aperture, with the direction of this effect in agreement with that proposed by the vowel aperture segmentation hypothesis. However, this effect was only found in experimental condition 1, first syllable repetition, linked to syllable offset detection. This finding raises questions as to the validity of the evidence supporting the vowel aperture segmentation hypothesis, as it has been suggested (Content et al, in press) that the operations of syllabic segmentation at the pre-lexical level of processing are driven by syllable onset detection.

Other findings of relevance to the significance of the segmentation effect of vowel aperture are those concerning repetition errors involving vowel aperture, referred to as vowel aperture errors. These describe subject responses where there was disparity between the vowel aperture produced in the response, and that of the stimuli. These errors were found in 27% of all of the subject responses in experimental condition 1. Two hypotheses were presented to explain the occurrence of these errors, the production and perception hypotheses. The first hypothesis suggests that subjects would tend to produce vowel aperture to match that of the syllable, with open vowels produced for closed syllables, and vice versa. The second hypothesis states that subjects are repeating the vowel aperture that was heard, and that an error was made in the perception of the vowel. Analyses revealed that suggested that there was evidence to support both of these hypotheses, suggesting that the mismatches between the vowel aperture used in the subject response and the stimuli were due to factors of both perception and production. However, whilst vowel aperture production errors have no bearing upon the segmentation results, the errors in perception reduce the possible effects of vowel aperture upon syllable segmentation.

This finding raises an important question regarding the degree of the segmentation effect due to vowel aperture. If subjects' segmentation responses are referenced using the aperture of the vowel found in the stimuli there is only a 9% increase in open syllable segmentation responses for close, rather than open vowels. However, if the vowel aperture is referenced by that produced by the subject, reducing the effect of vowel aperture perception errors, the effect of vowel aperture increases substantially, with a 30% increase in open syllable segmentation responses for close, rather than open vowels. Unfortunately, it is difficult to arrive at a decision as to which of these results should be used as the most accurate measure of the effect of vowel aperture upon segmentation. It is, perhaps, sufficient to state that vowel aperture has a statistically significant effect over syllable segmentation, with the direction of this effect in agreement with that predicted by the vowel aperture segmentation hypothesis.

The problems encountered with the vowel aperture errors in this experiment also highlight the concerns of Treiman and Danis (1988), in their examination of the effects of vowel length upon the segmentation responses of English experimental subjects. As discussed in Section 4.2, Treiman and Danis (1988) suggest that the differences between the segmentation of stimuli with short and long vowels could be coloured by production constraints. That is, when the first syllable contained a short vowel, subjects may have been more likely to close the syllable with a consonant, to produce a phonologically legal response, with the reverse case for long vowels. As the syllable repetition task used in Experiment 6 also relies upon the production of syllables in isolation, the same criticism may also be levelled at the results of this experiment. Therefore, to address these issues, and those arising from the disparity of the effect of vowel aperture between syllable onset and offset detection an additional experiment was required. By using an on-line task in Experiment 7 it was possible to ascertain whether the locus of the vowel aperture segmentation effect was post or pre-lexical, also subjects were not required to produce isolated syllables, and, as such, were unlikely to be affected by productional constraints.

In Experiment 7, a fragment detection task was used to ascertain whether the reaction times of the subjects would exhibit the 'syllable effect' arising from the match or mismatch of syllable boundaries between target and carrier stimuli. In this experiment syllable boundaries were referenced, not by the nature of the intervocalic consonant cluster/singleton, but by the aperture of the first vowel of the bi-syllabic stimuli. Syllable boundaries for carrier items were referenced according to the vowel aperture segmentation hypothesis. Carriers with an open first vowel were said to have a /VC.CV/ segmentation, whilst those with a close first vowel were said to have a

/V.CCV/ segmentation. The analyses of reaction times revealed a 'syllable effect' interaction between the syllable structure of the carrier and the target, with faster reaction times when the target matched the first syllable of the carrier than when there was a mismatch. This finding would appear to support the vowel aperture segmentation hypothesis, as reaction times were faster when the target fragment corresponded to the segmentation predicted by the hypothesis. Additional analyses also revealed that the locus of this effect was more likely to be at the pre-lexical level of processing rather than the post-lexical level of processing. If the segmentation effect of vowel aperture was based on a lexical or post-lexical level of processing it would be expected that the syllable effect would be strongest with the slower subjects, where the responses were of sufficient duration to allow a post-lexical access response. However, with a stronger syllable effect in the responses of faster subjects (whose reaction times would appear to eliminate the chances of a post-lexical access response), than the slower subjects, the possible use of metalinguistic knowledge would be appear to hamper the effect of vowel aperture upon syllabic segmentation, rather than inducing this effect.

In conclusion, the evidence of experimental condition 1 of experiment 6 and experiment 7 support the vowel aperture segmentation hypothesis, that vowel aperture is used as a syllable segmentation cue, with open vowels inducing closed syllables, and close vowels inducing open syllables. In addition, it would appear that this effect is more likely to be based upon the pre-lexical, rather than lexical or post-lexical processing level, which, in turn, supports the hypothesis that it is the syllable that is used as the pre-lexical processing unit in French.

7 Conclusions

7.1 Summary

Phonology is replete with theories and principles governing the process of syllable segmentation. However, with a plethora of conflicting theories available, a major problem facing researchers wishing to apply the syllable unit in models of speech processing is in selecting which of these theories to use in their definition of the syllable.

This study has concentrated upon the syllabification of the French language. This language has a number of inherent advantages for the study of syllable segmentation: it is a syllable-timed language, and is thought to have a relatively clear syllabic structure, at least, compared to stress-timed languages such as English. The syllable also has special, psycholinguistic, relevance in French, a language for which there is possibly the greatest collection of experimental evidence supporting the use of the syllable as a unit of pre-lexical processing.

In the study presented in this thesis, an attempt has been made to judge the merits of the various principles and theories of syllabification. By comparing the theoretical predictions of these theories with the experimental findings of a series of psycholinguistic syllable segmentation and perception experiments, it is hoped that some of the confusion and ambiguity surrounding the subject of syllable segmentation may be cleared, allowing syllable boundary placement to be predicted with greater accuracy. A number of factors have been found to influence syllable segmentation in French. A summary of the main findings of the various syllable segmentation and perception findings is covered below:

Syllable Segmentation Preference Laws based upon the regularities of Intervocalic Consonant Clusters and Singletons

In a comparison of a number of French-specific syllabification algorithms, it is the segmentation preference laws of Laporte that appear to be closest to French segmentation behaviour. These laws are based upon the phonotactic regularities of intervocalic consonant clusters and singletons, predicting that the syllable boundary is placed before the last segment which is not a glide, with a subset of OBLI's ([p t k b d g f v] followed by [l r]) treated as a single segment. However, it is suggested that the clusters /tl/ and /dl/ be excluded from the set of OBLI's defined by Laporte, to fit with the definition of these clusters suggested by Dell.

In essence, these rules form that of a Minimum Allowable Onset Principle, that the syllable boundary be placed such that the onset is of minimum length so long as it is in agreement with the Obligatory Onset Principle and the special treatment of OBLI clusters.

The Ambisyllabic Onset Hypothesis

To explain the differences in syllabification consistency between syllable segmentation via syllable onset and offset detection, the 'ambisyllabic onset hypothesis' was proposed. It states that if there is a legal syllable onset before the first consonant in an intervocalic consonant cluster/singleton then there will be a high degree of increased onset ambisyllabicity for that consonant. This predicts higher syllabification consistency for syllable onset detection than offset detection when the syllable boundary of the preferential segmentation lies before the first consonant, with the disparity reversed if the preferential response lies behind the consonant.

Vowel Aperture Segmentation Hypothesis

The vowel aperture segmentation hypothesis states that if the vowel of a syllable is open then this will induce syllable segmentation responses that result in the syllable being closed, if the vowel is close this will induce segmentation responses that result in an open syllable. Two experiments, one using a metalinguistic task, the other an online task, were used to investigate the claims made by the vowel aperture segmentation hypothesis. Evidence from the subject responses in both experiments appears to support the hypothesis, suggesting that vowel aperture, as well as the phonotactic regularities of the intervocalic consonant cluster/singleton, is used as a cue in French syllabic segmentation. It was also found that the degree of variation in subject segmentation responses due to vowel aperture was dependent upon the overall variability found in segmentation responses for a particular intervocalic consonant cluster or singleton, as predicted by the ambisyllabic onset hypothesis. Analyses also reveal that the locus of the vowel aperture segmentation effect is likely to be at the pre-lexical, rather than the lexical or post-lexical level of processing. It is also suggested that this finding supports the theory that the syllable is the pre-lexical processing unit for the French language.

The role of Orthography

With the onset of literacy, another potential cue that may be used in the segmentation decisions of subjects in metalinguistic experiments are those resulting from orthographic bias. To investigate the scope of orthographic bias upon the responses of experimental subjects, the syllabification behaviour of two subject groups were compared, those of pre-literate children, and literate adults. With knowledge of syllabic structure, but no knowledge of orthography, one difference in the syllable segmentation behaviour of pre-literate children and literate adults is due to the acquisition of orthographic knowledge. A comparison of the segmentation responses of the two subject groups revealed that the only differences in the syllabification behaviour of these two subject groups were limited to stimuli containing the /ks/ cluster. Therefore, it appears that the influence of orthographic bias upon syllable segmentation is limited to a few special cases, such as the cluster /ks/, and has no effect on the general syllabification strategies of listeners.

These findings suggest that listeners are taking advantage of a number of separate cues when segmenting speech into syllables. This is at odds with the majority of the phonological theories concerning syllabic segmentation, which base their predictions upon only a single cue, the phonotactic nature of the intervocalic consonant cluster. In this study only one additional syllabification cue has been examined, that of vowel aperture, which has been found to have a significant effect on the segmentation decisions of the experimental subjects. However, even when factoring the effect of vowel aperture into the segmentation responses of subjects, there is still considerable variability in subject segmentation responses. This would suggest that there might be other segmentation cues, as yet unexplored, which influence the syllable segmentation decisions of listeners. In light of these findings, it is not possible to state that the segmentation strategies employed by listeners can be predicted with complete accuracy. However, by implementing the factors found to influence syllabification thus far, it is possible to suggest a set of preference rules which can predict where the boundary is *likely* to be located, and also to predict in which situations the location of the syllable boundary is likely to be most ambiguous.

Therefore, in closing this thesis, a set of syllabification preference rules are presented, based upon both the findings of the syllable segmentation and perception experiments performed in this study, and an examination of phonological syllabification theory:

If there only a single legal syllable onset for a particular intervocalic consonant or singleton, and if this onset is not adjacent to a vowel, there is a strong preference for the placement of the syllable boundary before the first segment of this onset. In all other cases the location of the syllable boundary can be considered as ambiguous, with syllable segmentation decided using three additive segmentation preference rules.

- *If there is a legal syllable onset before the first consonant in an intervocalic consonant cluster/singleton then there is a preference for the placement of the syllable boundary before that consonant for syllable onset detection, and after for syllable offset detection.*

- *There is a preference for the placement of the syllable boundary before the last segment which is not a glide, with a subset of OBLI's ([p t k b d g f v] followed by [l r]), excluding clusters /tl/ and /dl/, treated as a single segment.*
- *If the vowel preceding the intervocalic consonant cluster/singleton in question is close there is a preference for segmentation resulting in an open first syllable, if the vowel is open there is a preference for segmentation resulting in a closed first syllable.*

If all of these preferential segmentation rules agree as to the placement of the syllable boundary, then it is expected that syllabification consistency will be high. However, if there are conflicting syllable segmentation preferences then there will be greater ambiguity surrounding the placement of the syllable boundary.

7.2 General Discussion

7.2.1 Syllable Segmentation, Perception and Acquisition

As has been described in the summary, the most influential components of syllabification behaviour in French listeners, typified by the behaviour described by their preferential syllabification responses, can be modelled using a set of relatively simple rules, based upon the application of concepts of legality.

When exploring rules established from empirical data a classical question is that one must ask is how these rules are implemented in the mind of the subjects, or, in other words, whether they have any psychological reality. As the information used in speech perception lie in the knowledge of the listener, and not the signal, it is the human mind, not the acoustics of speech that must be the primary focus of any investigation. Whilst it must be kept in mind, as pointed out in at the start of this thesis, that phonological theory cannot be interpreted literally, as a direct description of mental representations, it is important to try and understand the bearing these rules might have on the processes of speech perception and acquisition.

It has been shown that, in adults, the rules used by listeners are highly language-specific and as a consequence the performance of the listeners will remain language-specific. This mirrors the findings found in the examination of pre-lexical speech segmentation, where subjects will export the segmentation behavior of their maternal language to other languages, even if this behavior is not well suited to the processing of these new languages (Cutler, Mehler, Norris, &

Segui, 1986; Dupoux, Kakehi, Hirose, Pallier, & Mehler, 1999) (although this could be changed through extensive training in some cases).

In language acquisition, the ability to acquire these language-specific rules has to be questioned. In this case, both the signal and the infant's learning mechanisms are required if segmentation rules are to be recovered from speech, that is, information contained in the signal will trigger the learning device. When and how are these rules recovered?

It has been suggested that broad syllabification strategies are acquired in the early stages of language development, most likely before the end of the first year. Evidence for this hypothesis comes from recent studies on infants' sensitivity to the rhythmic structure of their maternal language. It has been shown that, from birth, infants are able to classify the incoming languages as a function of their rhythmic class (moraic vs. syllable-timed vs. stress-timed), and that they are able to refine their categorisation within a given rhythmic class (differentiating amongst syllable-times languages for instance) (Nazzi, Bertoncini & Mehler, 1998; Christophe & Morton, 1998). Hence, selecting the segmentation strategy that is best adapted to a target language may allow children to specify the segmentation unit (which would be the syllable in syllable-timed languages for instance).

A second source of evidence comes from findings that infants have acquired some of the language-specific phonotactic regularities from the age of nine months. More precisely, infants of this age show greater familiarity with utterances that start and end with a legal phonemic sequence in their language. This does not indicate that infants learned these regularities by segmenting continuous speech into words, retrieving word boundary cues, but simply by listening to the very beginning and end of utterances. The indication that infants are aware of these regularities, just prior to the construction of their first perceptual mental lexicon, would suggest that they are *able* to use these cues to segment continuous speech.

When examining the mechanisms used to acquire syllabification strategies during language development another problem is evident, that of syllabification bootstrapping. The problem of bootstrapping describes the paradox that to syllabify a string you need to know syllabification rules, and to acquire the rules you need to know the rules in order to syllabify initial strings. Beside the traditional split between parameter setting theories and statistical learning, it is also possible that a relatively simple heuristic *à la* Slobin (Slobin, 1985(i) & Slobin, 1985(ii)) would explain how learning takes place. As mentioned above, if infants were to pay attention to the beginning and end of utterances, they could discover the Legality Principle of syllabification. This principle is the basis of the formation of the Laporte algorithm, when the principle is applied to French. This principle would lead to the extraction of language-specific rules, for it would leads to strategies that are language specific. This kind of proposition is very similar to that of Dahan and Brent (1999) who were attempting to circumvent the lexical bootstrapping problem. They suppose that children store relatively short utterances and code them as new, old, or made up of new and old parts, depending on what they have already encountered. With this rather simple algorithm an important set of lexical items can be recovered, and further learning will refine the stored entries by gradually specifying phonological representations. By applying a similar strategy to the acquisition of the syllabification mechanism it could be suggested that the Legality Principle is used as the foundation for French syllabification. Whilst the application of the Legality Principle cannot account for the entire syllabification strategy used in adults, it is a simple mechanism that can be used for bootstrapping, that can later be refined with additional, learned, syllabification cues, such as vowel aperture. It remains to determine whether this learning strategy can be applied to the syllabification of other languages for the formation of a universal syllable segmentation bootstrapping process.

7.2.2 Syllable Segmentation: boundaries, onsets or offsets?

Although it is seldom commented upon, it is commonly held that listeners syllabify speech by locating a boundary between two syllables. This is not to say that there is an explicit acoustic 'boundary' cue, but that it is decided from the broad context of information surrounding a potential boundary (such as nature of syllable nuclei and intervocalic consonants). One problem with the *boundary hypothesis* of syllabification is the notion of ambisyllabicity. Ambisyllabicity is incompatible with this hypothesis as it implies that the offset of one syllable does not coincide with the onset of the next, and thus, does not form a boundary. This incompatibility has been commented upon by Selkirk (1982) who stated that "Clearly, ambisyllabicitiy, if it exists, would provide further argument against the boundary approach to the syllable: a syllable boundary cannot be simultaneously before and after some segment of the string".

An alternative hypothesis, that proposes an explanation for ambisyllabic behaviour is the *syllable onset segmentation hypothesis* (Content et al, in press). This hypothesis makes a distinction between two different operations, syllable onset and offset detection, and emphases the importance of syllable onsets, suggesting that syllable segmentation is performed through the detection of syllable onsets. This hypothesis has similarities with the Legality Principle, with the split between the Law of Initials (dealing with onsets) and the Law of Finals (dealing with offsets), and the dominance of the Law of Initials over that of Finals, embodied in the Principle of the Irregular Coda, and taken still further with the Maximum Onset Principle. Content et al (in press) presented evidence to support this hypothesis by showing that syllabification consistency was higher in syllable onset, rather than offset detection. This was demonstrated using a syllable repetition task with single intervocalic consonant stimuli.

In this thesis the effect described by Content et al. (in press) was also found in Experiment 1. However, this effect was only consistent in single intervocalic consonant stimuli, and not with intervocalic consonant clusters, where the differences between syllable onset and offset consistency were dependent upon the intervocalic consonant cluster. To explain these findings a different hypothesis was proposed by the author, the *ambisyllabic onset hypothesis.* This hypothesis does not support the theory that the syllable onset holds special status in syllable segmentation, as it predicts that syllabification consistency can be higher in syllable onset detection than offset detection, *or vice versa,* depending on the preferential segmentation used by the listener.

Other evidence, from Experiments 6 and 7, also shed doubts on the special status of syllable onset detection. In Experiment 6 the vowel aperture segmentation effect was only found in experimental condition 1, associated with syllable offset detection. If syllable segmentation is performed by means of syllable onset detection then the effect of vowel aperture found in this experiment would have no relevance to normal syllable segmentation, as differences in vowel aperture had no effect in experimental condition 2, associated with syllable onset detection. In Experiment 7 the effect of vowel aperture was re-tested using a different experimental paradigm, the fragment detection task, a task designed to reveal syllable segmentation behaviour that makes no distinction between onset and offset detection. The results of this experiment revealed a significant effect of vowel aperture upon the syllable segmentation behaviour of listeners. If the pre-lexical segmentation of syllables is by means of onset detection, as proposed by Content et al. (in press) then why does the effect of vowel aperture, only found in syllable offset condition of Experiment 6, serve as a cue for the ‘syllable effect’ when tested with an on-line task? It is suggested that, because of the findings of experiments 6 and 7, and also from experiment 1, that it is unlikely that syllable segmentation is performed through the simple detection of syllable onsets.

The evidence found in this thesis would appear to point towards a segmentation strategy that takes advantage of multiple cues, found in syllable nuclei, offsets *and* onsets. It is possible that ambisyllabic behaviour is caused by experimental tasks that attempt to separate the processes of syllable onset and offset behaviour, such as the syllable repetition task. As there are conflicts inherent in the multiple cues used in syllable segmentation (such as those provided by the vowel aperture and the intervocalic consonant or cluster) it is possible that, by concentrating the listener on the syllable onset *or* offset that the segmentation behaviour of listeners could change, as they are more likely to be influenced by cues that are local to these areas of the syllable.

7.2.3 The Syllable and Artificial Speech Recognition

7.2.3.1 The 'State of the Art' in Artificial Speech Recognition

Until quite recently, there has been rapid development of Automatic Speech Recognition (ASR) systems, with word-error rates dropping by a factor of five between 1990 and 1995. In small recognition corpora, with a closed vocabulary (i.e. all words tested with the recogniser are contained within its lexicon), such as the TU connected digits corpus, can be recognised with a word error rate of only 0.3%. For larger, continuous speech applications, trained using the ARPA Wall Street Journal Dictation corpora (20,000 words) error rates of 13% for closed vocabulary, and 26% for open vocabulary have been achieved (Schwartz & Schwartz, 1995).

However, these recognition rates are only valid if you assign a double meaning to the 'artificial' of ASR, that is, they are designed to recognise speech under 'artificial' or unnatural conditions. Once you take these systems out of laboratory conditions and introduce them to 'real-world' environments recognition rates decrease dramatically. Noise and reverberation, commonplace elements in most acoustic environments, are particularly damaging to ASR performance, typically reducing word-level accuracy to 20-50% (Gong, 1995). For example, an isolated word recogniser that gives 100% accuracy with clean speech, can typically drop to 30% when used in a car travelling at 90 kmh (Lockwood & Boudy, 1991). In addition to the problems caused by their acoustic environment, current ASR technology also encounters grave difficulties when faced with naturally spoken discourse. Most recognisers are trained using 'read speech' that is, the recording of speakers reading from a text in a measured and controlled manner. When presented with 'spontaneous' speech, found in most natural environments, the diversity of speaking styles and 'sloppy' speech will increase word-error rates to around 50% (Cohen, Gish & Flanagan, 1994). Therefore, bring the latest 'state of the art' recogniser out of the laboratory and into a normal human environment, with it's inherent variability, and transcription will be reduced from a comprehensible facsimile of the spoken discourse to a stream of gibberish.

Because of these severe limitations there is growing feeling that the traditional theoretical framework for ASR is limiting performance, that the current technology is stuck in a 'local maxima' of performance, and will require radical changes in order to reap benefits in the long term (Greenberg, 1996; Bourlard, 1996).

7.2.3.2 The Syllable as the acoustic unit of recognition

Early in the history of artificial speech recognition the main area of research was related to the problem of how to quantise the rather complex waveform of speech into a simpler set of speech vectors for use in the recognisers of the time. To this end researchers turned to two of the popular speech analysis 'tools' of the time, the Sonograph and the phoneme.

The Sonograph allowed sound to be split into its constituent frequencies, making speech analysis a far simpler since the complex speech waveform can be split into less complex components for a more detailed view. The second 'tool', the phoneme, a set of mutually exclusive elementary units related to the articulatory gestures of a language. These possess a property such that if one phoneme in a word is replaced with another its meaning will change. Using these techniques speech researchers could now transform the speech waveform into standard sized vectors representing frequency powers over a set time 'frame'. Recognition could now be simplified to a process that could find the appropriate transform to match each time/frequency vector with the appropriate phonemic 'atom'.

It is the search for the 'perfect' transform between frame and phoneme that has dominated ASR research, from the very first phoneme recogniser (Dudley & Balashek, 1958) up until the present day. Whilst there have been modifications to these representations, using cepstral frequency representations, and the use of tri-phone instead of mono-phone acoustic models, the basic model is largely unchanged.

In response to the calls for radical changes in ASR technology, and an 'increase in error rates' there has been recent interest in a re-examination of the basic ASR model and how it could be made more robust to 'normal' working conditions. Both components of the basic ASR model are under examination, with a strong movement to make ASR more psychologically plausible, in other words, to try and model human speech perception mechanisms instead of driving for an artificial engineering solution. Firstly, there is interest in replacing the cross frequency spectral frame with independently processed frequency channels, prompted by psychoacoustic evidence (e.g. Allen, 1994; Warren, Riener, Bashford & Brubaker, 1995). Secondly, replacing the phoneme as the acoustic unit of recognition with the syllable, prompted by psycholinguistic evidence (such as that covered in chapter 1). Naturally, it is the second aspect that is of most interest to this study.

Using the syllable as the acoustic unit of recognition is not a new concept, and was under investigation as far back as 1975 (Fujimura, 1975), but has only recently has there been widespread study of this unit. There have been a number of reasons put forward for the use of the syllable in ASR, most importantly that the effects of coarticulation, the biggest problem in phoneme recognition (only partially modelled by the use of tri-phone models), are greater for phoneme boundaries than syllable boundaries. Psycholinguistic evidence, proposing that the syllable is used as a pre-lexical unit of speech perception, is also widely cited. Unfortunately, the controversy surrounding this subject in psycholinguistics, especially the differences found between languages, does not seem to have caught the attention of most speech technologists. As a typical example, in writing an introduction explaining the rationale behind the use of syllable boundaries in ASR (Wu, et al, 1997) the authors have stated that psycholinguistic evidence suggests that the syllable is a "basic perceptual unit for speech processing in humans". In supporting this statement they cite the study of Segui et al, (1990). What seems to have been overlooked is that psycholinguistic evidence supporting the pre-lexical use of the syllable is largely limited to syllable timed languages, and that in English (the language under analysis in the study of Wu et al, 1997) the majority of evidence does not support the syllable.

Overlooking the rather naïve attitude of most speech engineers concerning psycholinguistic evidence for the use of the syllable, there are also considerable engineering challenges involved in replacing the phoneme with the syllable in ASR. The most obvious concerns the number of units required in the acoustic model when using phonemes and syllables. Taking French as an example language, there are usually 39 phones modelled in a phonemic recogniser, if the syllable were used then at least the 3246 syllables found in a search of the BRULEX lexicon (Goldman et al, 1996) would have to be modelled. The increase in complexity from a phonemic to syllabic model is enormous, and would require a correspondingly massive increase in the amount of training data required for a syllabic model. Even if such a feat were possible with today's technology it would not be possible to simply replace the phonemic acoustic model and leave the rest of the ASR system intact. As has been described, current ASR is based upon the matching of a spectral slice of fixed length (usually around 20ms) with a phonemic model. Unfortunately this approach would not translate to syllable recognition, as the duration of syllables can vary from 20ms to around 300ms. Most ASR research has concentrated on the development of statistical methods (Hidden Markov Models and Artificial Neural Networks) specifically for the task of matching spectral slices on phonemes. Therefore, the speech engineer wishing to develop a pure syllabic large vocabulary recogniser (treating the syllable as an *atomic* unit) would be forced to abandon the vast majority of the techniques used in ASR for the past 40 years, and would, in effect, have to start from scratch.

Unsurprisingly research into the use of the syllable in ASR has begun rather tentatively, and has taken two main forms, the development of restricted vocabulary syllable based ASR, and the use of hybrid solutions, where syllabic information is used to enhance standard phonemic recognisers.

In the former case, the restricted vocabulary ASR is usually *highly* restrictive; typical studies of syllabic based ASR are limited to small corpora, such as OGI digits (Cole, Noel, Lander and Durham, 1995), consisting of only 92 different words. Two examples of syllable-based digit recognition show marginal improvements between syllabic and phonemic recognisers. In a study by Hauenstein (1997) the corpus was modelled using 96 individual syllables, and achieved a frame error rate of 16.2% compared to 17.3% for a similar phoneme based system. More impressively, when restricted to an 11-word digits corpus the error rate for the phoneme based system dropped from 4.5 to 25.1% after adding “car noise” with 10 dB SNR (Signal to Noise Ratio). The syllable based system dropped from 2.5% to 20.7%, showing that it was more robust to noise. In the second example, by Hamaker, et al (1998), a syllable-based system, modelling 42 distinct syllables, was compared to a tri-phone system, using 3225 tri-phones. A comparison of word error rates for the two systems found that there was marginally better performance for the syllable based system, with and error rate of 11.1% compared to 12.2% for the tri-phone system.

Similar improvements have also been noted for larger vocabularies, in a study by Jones, Downey and Mason (1997) a syllabic recogniser was trained on a 1243 word corpus, consisting of 1313 distinct syllables. In this study a syllabic recogniser was compared with a standard phonemic version using the same training and test sets. It was shown that syllable recognition accuracy approached 50%, improving on a phonemic recognition accuracy of only 36%. Interestingly the authors also noted that most recognition problems occurred due to the lack of training data for low frequency syllables. They showed that 95% of the individual syllables with poor (less than 5%) recognition accuracy had less than 30 training examples.

As well as ‘pure’ syllabic recognisers, hybrid approaches, where standard phonemic recognisers are enhanced with syllabic information, also show some success. In a study by Dupont, Bourlard and Ris (1997) multi-stream recognition of the OGI digits corpus was performed using two parallel models. In the first model the syllable is represented by concatenating phone models to capture the fine structure of the syllable. In the second model a longer temporal context is used to capture the gross temporal structure of the syllable. This hybrid system showed lower word error rates than a standard phonemic recogniser both in clean speech (8.9% vs. 10.7%) and with the addition of additive gaussian white noise at 15dB SNR (16.2% vs. 17.2%). In a collection of publications headed by Wu (Wu et al, 1997; Wu et al, 1998(i); Wu et al, 1998(ii)) a variety of syllabic cues have been added to standard phonemic recognisers. In one system a syllabic representation is inserted between the phoneme and word representations (Wu, et al, 1997). Initially phonemes are recognised, and, in addition, syllable onsets are also detected, the syllable onsets are used to group phonemes into syllables, which are then matched with words. In an approach similar to that of Dupont et al (1997), Wu et al (1998(i) & 1998(ii)) use two parallel streams, the first for the recognition of phonemes, the second for the recognition of syllables from low frequency modulation spectrograms. In all of these studies a marginal improvement was reported for recognition accuracy of the OGI digits corpus when compared to standard phonemic recognition.

7.2.3.3 Syllable Segmentation and Syllable based ASR

There are two essential requirements common to all ASR systems, a description of the acoustic representation of speech, represented as a list of all possible acoustic units or features possible in a particular language, and a lexicon that describes how the acoustic representation maps onto words. In the standard phonemic ASR system the fulfilment of these two requirements is not especially problematic. The first requirement, a listing of all broad phonemic categories for a particular language, is a central feature of phonetic theory and this information is widely available. Again, phoneticians have furnished the second requirement of a phonemic recogniser, large vocabulary phonetically labelled lexica are now commonplace.

When using syllabic information for ASR similar, syllabic, equivalents are necessary, a list of syllables that composes a language, known as a syllabary, and a lexicon with syllable markings. However, as the syllable enjoys less currency than the phoneme in speech research, and perhaps due to the controversy surrounding syllabic segmentation, it is far more difficult to fulfil these requirements. There are relatively few lexica containing syllabic markings, and still fewer published syllabaries. Therefore, it is usually necessary to turn to syllabification algorithms to construct these resources, and thus, the researcher wishing to perform this task is faced with the question that has formed the basis of this thesis, which syllabification strategy do you use? This problem has largely been overlooked in ASR research. Of all the investigations reviewed in this section only one gave mention of the mechanisms that were used in defining the syllable (Jones et al. (1997) used the Maximum Onset Principle to syllabify their 1243 word corpus). Admittedly, as has been seen, most investigations have been restricted to very small vocabularies, of around 90 words, where syllabary construction and segmentation can be performed ‘by hand’. However, if progress is to be made then it is clear that vocabularies will need to be extended. This will require the use of an accurate mechanism of syllable segmentation, for the construction of syllabaries and the marking of lexica. Without an accurate means of syllable segmentation, meaning a mechanism corresponding to human perception and segmentation strategies, then the vaunted advantages of the syllable over the phoneme for ASR will be diluted, if not lost. To take advantage of the reduced coarticulation between syllable boundaries you need to *find* those boundaries, if mistakes are made then the advantage is lost. Instead of syllabic recognition you have the recognition of arbitrary groupings of vowels and consonants, which are unlikely to bring the breakthrough in robust speech recognition that has been proposed for the use of the syllable.

Bibliography

Alegria, J., & Morais, J. (1979). The development of the ability of conscious phonetic analysis of speech and the learning of reading. *Archives de Psychologie, 47(183),* pp.251-270.

Barnwell, T. P. (1970). An algorithm for segment duration in a reading machine context. Unpublished doctoral thesis, Massachusetts Institute of Technology.

Bertoncini, J., Floccia, C., Nazzi, T. & Mehler, J. (1995). Morae and syllables: Rhythmical basis of speech representations in neonates. *Language and Speech, 38*, pp.311-329.

Bertoncini, J. & Mehler J. (1981). Syllables as units in infant speech perception. *Infant Behavior and Development, 4*, pp. 247-260.

Bijeljac-Babic, R., Bertoncini, J. & Mehler, J. (1993). How do 4-day-old infants categorize multisyllabic utterances? *Developmental Psychology, 29*, pp.711-721.

Bourlard, H.A., Hermansky, H., & Morgan, N. (1996). Towards increasing speech recognition error rates. *Speech Communication, 18(3)*, pp.205-231.

Bradley D.C., Sànchez-Casas R.M., & Garcìa-Albea J.E. (1993). The status of the syllable in the perception of Spanish and English. *Language and Cognitive Processes, 8,* pp.197-233.

Calmès, M. & Pérennou, G. (1998). BDLex : a Lexicon for Spoken and Written French. *Proceedings of the 1st International Conference on Langage Resources & Evaluation*, pp.1129-1136.

Chomsky, N., & Halle, M. (1968). The sound pattern in English. New York: Harper & Row.

Christophe, A., & Morton, J. (1998). Is Dutch native English? Linguistic analysis by 2-month-olds. *Developmental Science, 1*, pp.215-219.

Church, K. W. (1983). Phase structure parsing: a method for taking advantage of allophonic constraints. Unpublished doctoral thesis, Massachusetts Institute of Technology.

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. J. Kingston & M. Beckman (Eds.), Between the grammar and physics of speech, New York: Cambridge University Press, pp. 283-333.

Cohen, J. Gish, H. & Flanagan, J. (1994). Switchboard - the second year, *CAIP Summer Workshop in Speech Recognition: Frontiers in Speech Processing II, July 1994.*

Cole, R.A., Noel, M., Lander, T., & Durham, T. (1995). New telephone speech corpora at CSLU. *Proceedings of Eurospeech 1995,* pp.821-824.

Coltheart, M. (1978). Lexical access in simple reading tasks. G.Underwood (Ed.), Strategies of Information Processing, New York: Academic Press.

Content, A., Mousty, P. & Radeau, M. (1990). BRULEX: Une base de données lexicales informatisée pour le français écrit et parlé. Manuscript of the Université Libre de Bruxelles.

Content, A., Kearns, R. K. & Frauenfelder, U. (in press). Boundaries versus onsets in syllabic segmentation. *Journal of Memory and Language.*

Coursil, J. (1992). Grammaire Analytique du Français Contemporain. Essai d'Intelligence Artificielle et de Linguistique Générale. Unpublished doctoral dissertation, Université de Caen.

Cutler, A., Mehler, J., Norris, D. & Segui, J. (1983). A language-specific comprehension strategy. *Nature, 304*, pp.159-160.

Cutler, A., Mehler, J., Norris, D., Segui, J. (1986). The syllables differing role in the segmentation of French and English, *Journal of Memory and Language, 25*, pp.385-400.

Cutler, A. & Otake, T. (1994). Mora or phoneme ? Further evidence for language-specific listening. *Journal of Memory and Language, 33,* pp.824-844.

Dahan, D., & Brent, M. R. (1999). On the discovery of novel wordlike units from utterances : An artificial-language study with implications for native-language acquisition. *Journal of Experimental Psychology: General, 128(2),* pp.165-185.

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

Dudley, H. & Balashek, S. (1958). Automatic Recognition of Phonetic Patterns in Speech*, Journal of the Acoustical Society of America, 30,* pp.731-733.

Dumay, N. & Radeau, M. (1997). Rime and syllabic priming effects in phonological priming between French spoken words. Proceedings of the *5th European Conference on Speech Communication and Technology, Rhodes, Greece*.

Dupont. S, Bourland, H. & Ris. C. (1997). Robust speech recognition based on multi-stream features. *Proceedings of the ESCA-NATO Workshop on Robust speech recognition for unknown communication channels, 1997.* pp.95-98.

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),* pp.1568-1578.

Eimas, P.D., Siqueland, E. R., Jusczyk, P.W. & Vigorito, J. (1971). Speech perception in infants. *Science, 171,* pp.303-306.

Fallows, D. (1981). Experimental evidence for English syllabification and syllable structure. *Journal of Linguistics, 17*, pp.309-317.

Ferrand, L., Segui, J. & Grainger, J. (1996). Masked priming of word and picture naming : the role of syllabic units. *Journal of Memory and Language, 35,* pp.708-723.

Ferrand, L., Segui, J. & Humphreys, G. W. (1997). The syllable's role in word naming. *Memory and Cognition, 25,* pp.458-470.

Floccia, C. & Goslin, J. (submitted for publication). Syllabification rules for intervocalic consonant clusters in French: a developmental perspective, *Submitted to the Journal of Child Language*.

Friederici, A. D. & Wessels, J. M. I. (1993). Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception and Psychophysics, 54*, pp.287-295.

Fujimura, O. (1975). Syllable as a unit of speech recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 23(1),* pp.82-87.

Fujimura, O. & Lovins, J. (1982). Syllables as concatenative phonetic units, Indiana University Linguistics Club.

Gillis, S. & De Schutter, G. (1996). Intuitive syllabification: Universals and language specific constraints. *Journal of Child Language, 23*, pp.487-514.

Goldman, J.-P., Content, A. & Frauenfelder, U. H. (1996). Comparaison des structures syllabiques en français et en anglais. *Proceedings of the XXIst Journées d'Etude sur la Parole, Avignon,* pp.119-122.

Gong, Y. (1995). Speech recognition in noisy environments: a survey. *Speech Communication, 16,* pp.261-291.

Goswami, U., & Bryant, P. (1990). *Phonological Skills and Learning to Read.* Hillsdale, N.J.: Lawrence Erlbaum.

Greenberg S. (1996). Understanding speech understanding: Towards a unified theory of speech perception. *Proceedings of the ESCA Workshop on the Auditory Basis of Speech Perception, Keele, UK, July 1996.* pp.1-8.

Grosjean, F. & Gee J.P. (1987). Prosodic structure and spoken word recognition. *Cognition, 25,* pp.135-155.

Hamaker, J., Ganapathiraju, A., Picone, J. & Godfrey, J.J. (1998). Advances in alphadigit recognition using syllables. *Proceedings of the International Conference of Acoustics and Speech Signal Processing, 1998.* pp. 421-424.

Hauenstein, A. (1997). Using syllables in a hybrid HMM-ANN recognition system. *Proceedings of Eurospeech, 1997.* pp. 1203-1206.

Hooper, J. B. (1972). The syllable in phonological theory. *Language, 48*, pp.525-540.

Jiang, T., & Peng, D. (1999). Chinese phonological awareness of children and the difference between good and poor readers. *Acta Psychologica Sinica, 31(1),*pp. 60-68.

Jones, R.J., Downey, S., & Mason, J.S. (1997). Continuous speech recognition using syllables. *Proceedings of Eurospeech, 1997.* pp.1171-1174.

Jusczyk, P.W., Pisoni, D.B. & Mullenix, J. (1992). Some consequences of stimulus variability on speech processing by 2-month-old infants. *Cognition, 43,* pp.253-291.

Kahn, D. (1976). Syllable based generalisations in English phonology. Bloomington, IN: Indiana University Linguistics Club.

Kessler, B.,Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language, 37,* pp.295-311.

Klatt, D.H. (1980). Speech perception: A model of acoustic-phonetic analysis and lexical access. R.A. Cole (Ed.), Perception and Production of Fluent Speech, Hillsdale, N.J.: Erlbaum.

Kohler, K. L. (1966). Is the syllable a phonetic universal? *Journal of Linguistics, 2,* pp.207-208.

Kolinsky, R. (1998). Spoken word recognition: a stage-processing approach to language differences. *European Journal of Cognitive Psychology, 10*, pp.1-40.

Kolinsky, R. & Morais, J. (1993). Intermediate representations in spoken word recognition : A cross-linguistic study of word illusions. *Proceedings of the 1993 Eurospeech Conference*, *Berlin*, pp.731-734.

Kolinsky, R. & Morais, J. (1997). Migrations in speech recognition, A guide to spoken word recognition paradigms, F. Grosjean and U. H. Frauenfelder (Eds.), East Sussex, UK : Psychology Press.

Kolinsky, R., Morais, J. & Cluytens, M.(1995). Intermediate representations in spoken word recognition : Evidence from word illusions. *Journal of Memory and Language, 34,* pp.19-40.

Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition, 38,* pp.245-294.

Laporte, E. (1993). Phonetic syllables in French: combinations, structure, and formal definitions. *Acta Linguistica Hungarica, 41*, pp.175-189.

Lebrun, Y. (1966). Sur la syllabe, sommet de sonorité. *Phonetica, 14,* pp.1-15.

Levelt, W. J. M. (1989). Speaking : from intention to articulation. Cambridge, Mass., MIT Press.

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

Liberman, I.Y., Shankweiler, D., Fischer, F.W. & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology, 18.* pp.202-212.

Liberman, A. M. & Studdert-Kennedy, M. (1978). Phonetic perception. H. Leibowitz & H. L. Teuber (Eds.), Handbook of Sensory Physiology: vol VIII: Perception, R. Held, Heidelberg: Springer.

Lleo, C. & Prinz, M. (1996). Consonant clusters in child phonology and the directionality of syllable structure assignment. *Journal of Child Language, 23,* pp.31-56.

Lockwood, P. & Boudy, J. (1991). Experiments with a Non-linear Spectral Subtractor, HMM's and the projection, for robust speech recognition in cars, *Proceedings of Eurospeech 1991, Vol 1.* pp.79-82

Marslen-Wilson, W. D. & Warren, P. (1994). Levels of perceptual representation and process in lexical access : Words, phonemes and features. *Psychological Review, 101,* pp.653-675.

Malmberg, B. (1963). Structural linguistics and human communication. An introduction into the mechanism of language and the methodology of linguistics. Berlin.

Schwartz, M. & Schwartz, R. (1995). State of the Art in Continuous Speech Recognition, *Proceedings of the National Academy of Science, USA, 92*, pp.9956-9963.

Marslen-Wilson, W. D. & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10,* pp.29-63.

Mehler, J. (1981). The role of syllables in speech processing: Infant and adult data. *Philosophical Transactions of the Royal Society, B 295,* pp.333-352.

Mehler, J., Dommergues, J., Frauenfelder, U. H., & Segui, J. (1981). The syllable’s role in speech segmentation. *Language and Cognitive Processes, 4,* pp.57-67.

Meunier, C., Frauenfelder, U., Content, A., & Kearns, R. (1997) The locus of the syllable effect: Prelexical or Lexical?, *Proceedings of the Fifth European Conference on Speech Communication and Technology, Rhodes.*

Moon, C., Bever, T. G., & Fifer, W. P. (1992). Canonical and non-canonical syllable discrimination by two-day-old infants. *Journal of Child Language, 19(1),* pp.1-17.

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

Morais, J., Kolinsky, R. & Nakamura, M. (1996). The psychological reality of speech units in Japanese. T. Otake and A. Cutler (Eds.), *Phonological structure and Language Processing : Cross-Linguistic Studies.* Berlin : Mouton de Gruyter.

Myers, S. (1987). Vowel shortening in English. *Natural Language and Linguistic Theory, 5,* pp.485-518.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance, 24(3),* pp.1-11.

Oller, D. K. & Eilers, R. (1975). Phonetic expectation and transcription validity. *Phonetica, 31,* pp.288-304.

Pallier, C., Sebastian-Gallés, N., Felguera, T., Christophe, A., & Mehler, J. (1993). Attentional Allocation within the syllabic structure of Spoken Words. *Journal of Memory and Language*, *32,* pp.373-389.

Pascual Leone, J., (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologica*, *32*, pp.301-345.

Peereman, R. (1998). Personal communication.

Pisoni, D.B. & Luce, P.A. (1987). Acoustic-Phonetic representation in word-recognition. *Cognition, 25,* pp.21-52.

Pitt, M. A., & Samuel, A. G. (1990). Attentional allocation during speech perception. How fine is the focus? *Journal of Memory and Language, 29*, pp.611-632.

Pulgram, E. (1970). Syllable, Word, Nexus, Cursus. The Hague: Mouton.

Radeau, M., Morais, J. & Segui, J. (1994). The effect of overlap position in phonological priming between spoken words. *Proceedings of the 1994 International Conference on Spoken Language Processing, vol 3*, pp.1419-1422.

Radeau, M., Morais, J. & Segui, J. (1995). Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology : Human Perception and Performance, 21,* pp.1297-1311.

Samuel, A.G. (1989). Insights from a failure of selective adaptation : Syllable-initial and syllable-final consonants are different. *Perception and Psychophysics, 45,* pp.485-493.

Saussure, F. de. (1916). Cours de linguistique générale. Lausanne and Paris: Payot.

Sebastiàn-Gallés, N., Dupoux, E., Segui, J., & Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language, 31,* pp.18-32.

Segui, J., Dupoux, E. & Mehler, J. (1990). The role of the syllable in speech segmentation, phoneme identification, and lexical access. G. Altmann (Ed.), *Cognitive Models of Speech Processing. Psycholinguistic and Computational Perspectives.* pp.263-280. Cambridge, Mass.: Bradford Books.

Selkirk, E. O. (1982). The syllable. The structure of phonological representations (Part II), H. Van der Hulst and N. Smith (Eds.), Dordrecht: Foris.

Slobin, D. I. (Ed.). (1985)(i). The crosslinguistic study of language acquisition: Vol. 1. The data. Hillsdale, NJ: Lawrence Erlbaum Associates.

Slobin, D. I. (Ed.). (1985)(ii). The crosslinguistic study of language acquisition: Vol. 2. Theoretical issues. Hillsdale, NJ: Lawrence Erlbaum Associates.

Slowiaczek, L. M. & Hamburger, M. B. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology : Learning, Memory and Cognition, 18,* pp.1239-1250.

Titone, D. & Connine, C. M. (1997). Syllabification strategies in spoken word processing: Evidence from phonological priming. *Psychological Research, 60*, pp.251-253.

Tranel, B. (1988). The sounds of French: an introduction. Cambridge University Press.

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

Treiman, R. & Zukowski, A. (1990). Toward an understanding of English syllabification. *Journal of Memory and Language, 29,* pp.66-85.

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

Trubetskoy, N. S. (1939 [1969 Translation]). Principles of Phonology. Berkeley: University of California Press.

Umeda, N. (1975). Vowel duration in American English. *Journal of the Acoustical Society of America, 68(2).*

Van Ooijen, B., Bertoncini, J., Sansavini, A. & Mehler, J. (1997). Do weak syllables count for newborns? *Journal of the Acoustical Society of America, 102(6),* pp.3735-3741.

Vennemann, T. (1988). Preference Laws for Syllable Structure and Explanation of Sound Change. Berlin: Mouton de Gruyter.

Wu, S-L., Kingsbury, B.E.D., Morgan, N. & Greenberg, S. (1998)(i). Incorporating information from syllable length time scales into automatic speech recognition. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, 1998.* pp.721-724.

Wu, S-L., Kingsbury, B.E.D., Morgan, N. & Greenberg, S. (1998)(ii). Performance improvements through combining phone and syllable-scale information in automatic speech recognition. *Proceedings of the International Conference on Spoken Language Processing, 1998,* pp. 459-462.

Wu, S-L., Shire, M.L, Greenberg, S, Morgan, N. (1997). Integrating syllable boundary information into speech recognition, *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, 1997.* pp.987-990.

Zwitserlood, P., Schriefers, H., Lahiri, A. & van Donselaar, W. (1993). The role of syllables in the perception of spoken Dutch. *Journal of Experimental Psychology : Learning, Memory and Cognition, 19,* pp.260-271.

Appendix A: Stimuli used in Experiment 1

F	L	N	P	FF	FL	FN	FP	LF	LL	LN	LP
/usy/	/ula/	/una/	/yta/	/ysfi/	/ivru/	/iznu/	/usty/	/arʃy/	/urlu/	/yrmu/	/yrdu/
/uzy/	/iri/	/imi/	/ipa/	/isfa/	/izla/	/azmu/	/ifty/	/yrva/	/arla/	/urni/	/yrbi/
/avu/	/aly/	/any/	/udu/	/asfy/	/avra/	/uvny/	/azby/	/irzu/	/yrlu/	/ilma/	/yldu/
/usi/	/yra/	/yma/	/ity/	/ysfu/	/azlu/	/yzna/	/istu/	/irʃu/	/irly/	/armu/	/irda/
/yza/	/ulu/	/unu/	/apu/	/usfa/	/uvry/	/uzmy/	/ufta/	/arvy/	/irli/	/yrny/	/urby/
/iva/	/iru/	/uma/	/idy/	/ysfu/	/uzlu/	/avnu/	/yzba/	/urza/	/yrly/	/ulmi/	/ildy/
NF	**NL**	**NN**	**NP**	**PF**	**PL**	**PN**	**PP**	**LPL**	**NLP**	**PFP**	
/umvu/	/inly/	/umni/	/indu/	/aksu/	/ugru/	/ugmy/	/akti/	/yrblu/	/ungli/	/akspu/	
/anvy/	/ymla/	/unmi/	/ambi/	/ygzu/	/ybry/	/idmi/	/upty/	/urbli/	/angla/	/yksti/	
/amvy/	/imry/	/amny/	/angu/	/ypsu/	/ugla/	/ubma/	/abdi/	/irbla/	/ingly/	/ukspa/	
/unvu/	/anli/	/ynma/	/inda/	/yksa/	/agry/	/agmi/	/ikta/	/ardri/	/ingri/	/ikstu/	
/imva/	/umlu/	/imny/	/umby/	/igzy/	/ubri/	/ydma/	/ypti/	/irdra/	/ungra/	/ykspy/	
/ynvi/	/ymra/	/inma/	/yngu/	/upsi/	/igla/	/ibmu/	/ubdu/	/yrdru/	/ingru/	/aksta/	

Appendix B: Legal Onsets for Stimuli used in Experiment 1

Consonant Cluster	Number of Legal Onsets	Consonant Cluster	Number of Legal Onsets
/sf/	2	/nv/	1
/vr/	2	/nl/	1
/zl/	2	/ml/	1
/zn/	1	/mr/	1
/zm/	1	/mn/	2
/vn/	1	/nm/	1
/st/	2	/nd/	1
/ft/	2	/mb/	1
/zb/	1	/ng/	1
/rʃ/	1	/ks/	2
/rv/	1	/gz/	2
/rz/	1	/ps/	2
/rl/	1	/gr/	2
/rm/	1	/br/	2
/rn/	1	/gl/	2
/lm/	1	/gm/	1
/rd/	1	/dm/	1
/rb/	1	/bm/	1
/ld/	1	/kt/	2
/mv/	1	/pt/	2
		/bd/	1

Appendix C: Stimuli used in Experiment 2

LF	PF	LP	LN	PL	FP	FL	LL	PN
argent /rg/	absent /bs/	alcool /lk/	arme /rm/	abri /br/	basket /sk/	affreux /fr/	guirlande /rl/	admire /dm/
berceau /rs/	accent /ks/	altesse /lt/	armoire /rm/	adresse /dr/	biscotte /sk/	avril /vr/	horloge /rl/	magnum /gn/
cerceau /rs/	action /ks/	arbitre /rb/	calmer /lm/	boucler /kl/	biscuit /sk/	chevreuil /vr/	hurler /rl/	techno /kn/
chercher /rʒ/	atchoum /tʒ/	ardoise /rd/	carnet /rn/	caprice /pr/	casquette /sk/	coffret /fr/	parler /rl/	
colchique /lʒ/	boxeur /ks/	calcul /lk/	charmant /rm/	citron /tr/	castor /st/	couvrir /vr/		
courgette /rj/	capsule /ps/	cartable /rt/	cornet /rn/	eclair /kl/	espace /sp/	gaufrette /fr/	**PP**	**FN**
garcon /rs/	excite /ks/	corbeau /rb/	dernier /rn/	ecrire /kr/	espion /sp/	gifler /fl/	acteur /kt/	cosmos /sm/
marcher /rʒ/	gadget /dj/	culbute /lb/	dormir /rm/	endroit /dr/	espoir /sp/	givrer /vr/	capture /pt/	ovni /vn/
martien /rs/	klaxon /ks/	orchestre /rk/	filmer /lm/	epluche /pl/	histoire /st/	gonfler /fl/	dict̂e /kt/	
ourson /rs/	objet /bj/	pardon /rd/	fourmi /rm/	igloo /gl/	masquer /sk/	livret /vr/	docteur /kt/	
percer /rs/	observe /ps/	parking /rk/	journal /rn/	maigrir /gr/	past/que /st/	offrir /fr/	facteur /kt/	
percher /rʒ/	scotcher /tʒ/	serpent /rp/	marmite /rm/	montrer /tr/	poster /st/	ouvrir /vr/	lecture /kt/	
sergent /rj/	taxi /ks/	soldat /ld/	palmier /lm/	oubli /bl/	respire /sp/	refrain /fr/	septembre /pt/	
servir /rv/	vaccin /ks/	tartine /rt/	permis /rm/	remplir /pl/	rester /st/	ronfler /fl/		
urgence /rj/	vexer /ks/	tortue /rt/	tourner /rn/	tableau /bl/	risquer /sk/	siffler /fl/		
valser /ls/		volcan /lk/	vernis /rn/	tigrê /gr/	rosbeef /sb/	souffler /fl/		

Appendix D: Stimuli used in Experiment 3

FL	ʃr	ʃr	fl	fr	fr	jl	jl	jr	sl	sr	vl	vl	vr	zl	zl	zr	
	toʃral	øʃru	oflɛ̄	lyfrad	yfril	aʒlo	pyʒlɛ̄	uʒra	aslan	mysrɛ̄	ɛ̄vli	navlɔ̄	davriʃ	ɑ̄zle	nizlɑ̄	pazre	
FN	ʃm	ʃm	ʃn	sm	sm	sm	sm	vn	vn	vn	zm	zm	zm	zn	zn	zn	
	kɔʃmi	ɔ̄ʃmo	pɛʃnu	ysmɛ	usmir	dɔsmar	masmi	bovny	øvnal	ɛvnɑ̄	izmɑ̄	lozmɔr	fizmyt	riznom	aznym	oznap	
FP	ʃt	ʃt	ft	ft	ft	sk	sk	sk	sp	sp	zb	zb	zb	zb	zg	zg	
	iʃtɔn	eʃtal	poftar	piftyʃ	daftyl	yskɛn	lisku	toskin	syspav	ispœr	azbi	razbil	rɛ̄zbu	azbyr	ɔzgo	øzga	
LF	lʃ	lʃ	lf	lf	lj	ls	lv	lz	lz	rʃ	rf	rf	rj	rs	rv	rv	rz
	ɔlʃɑ̄	ɛlʃɑ̄	ilfɛ̄	nulfɛ̄	ylʒir	vɛlsɑ̄	sylvu	kalzɔʃ	alzɔ̄	mɔrʃɔt	pyrfi	urfe	ɑ̄rʒu	ʒarsik	ʒɛrvɑ̄	ɛrvir	yrzam
LL	lr	lr	rl	lr	rl	rl	lr	lr	lr	rl	rl	rl	lr	rl	lr	rl	rl
	fylrik	kalrɔn	ɑ̄rlyʃ	ylrø	dɔ̄rlo	barlin	ɛlra	tɔlrɛ̄t	ɑ̄lryʃ	ɔ̄rlɔs	marlɛt	œrlin	gulrav	ɔrluʃ	ilran	arlɔ̄	pirlɔ̄
LN	lm	lm	lm	lm	lm	lm	rm	rm	rm	rm	rm	rm	rn	rn	rn	rn	
	ylmik	kylma	pylmis	ulmœf	almɛ̄	dilmat	ɛrmɔ̄	fœrmak	ɔrmyʃ	pirmo	kyrmɛ̄	irmut	burne	arna	irnu	bœrnɔ̄	
LP	lb	ld	lg	lk	lk	lp	lt	lt	rb	rb	rd	rd	rg	rk	rp	rt	
	balbøp	ʃaldyt	ylge	mylki	olko	pɔlpɑ̄	sɔltɔ̄	kultɛ̄	urbyt	arbu	yrdɛg	lɛrdiv	irga	ɑ̄rkum	œrpis	tyrtɔ̄k	
PF	bz	bz	dv	dz	dz	dz	gz	gz	ks	ks	ks	ks	ps	ps	ps	pv	
	dabzɛr	ribzøt	idvuk	ydzal	adzɔʃ	ridzɔl	egzɛt	agzir	aksal	miksɔl	tikse	pɔksɑ̄	lypse	ɑ̄pse	mipsɔ̄	lapvik	
PL	bl	br	dr	dr	dr	gl	gl	kl	kl	kr	kr	pl	pr	tl	tl	tr	
	dɛ̄blu	obrɔ̄	gidrɔl	ydro	adrøt	øgla	pagli	teklar	ɛ̄kle	takrɔ̄	økre	oplɔg	dipra	etlu	vatlɛ̄	dytrif	
PP	bd	bd	kb	kd	kd	kt	kt	kt	kt	kt	pk	pk	pt	pt	pt	pt	
	ibdin	ɔ̄bdal	ɑ̄kbɛ̄	akdil	ukdan	vaktø	riktar	uktɔʃ	fyktan	dyktyl	apkul	ʃipkar	iptu	captin	liptɛr	ʃɑ̄ptɔ̄	
PN	bm	bm	dm	gn	gn	gn	km	kn	kn	pn	pn	tm	tm	tm	tn	tn	
	sybma	ɔbmy	admu	igna	tignu	magnyk	ykme	sykno	aknul	ypnir	tipne	fatmi	kɔtmɑ̄	itmuf	batny	rɛtnɔf	

Appendix E: Stimuli used in Experiment 5

Stimuli with Orthographic cues for first vowel aperture production							
Close Vowels				Open Vowels			
VC	Word	VC	Word	VC	Word	VC	Word
et	étude	om	chomer	ɔd	laideur	ɔr	foret
ot	autel	om	saumure	ɛd	soda	ɛs	aisselle
ot	autour	em	frémir	ɛt	laitance	ɔl	volant
ev	sévir	el	délice	ɔt	flottant	ɛt	bêta
ov	sauver	ol	frôler	ɛv	rêveur	ɛd	raidir
es	décu	or	aurore	ɔv	ovale	ɛt	traiter
os	aussi	en	bénir	ɛs	naissance	ɛt	vêtu
on	jaunir	er	chéri	ɔs	proces	ɔs	grossir
on	troner	ov	mauvais	ɛn	drainage	ɛn	peiner
				ɔn	tonnage	ɛm	aimer
				ɛm	aimant	ɛm	blêmir
				ɔm	fromage	ɛr	mairie
				ɛr	errant	ɛn	gênant

Stimuli with no orthographic cues for first vowel aperture production	
VC (BRULEX reference)	Word
ø	bleuter
ø	foetus
ø	jeûner
ø	neurone
ø	jeudi
œ	breuvage
œ	pleureur
œ	jeunesse
œ	gueulante

Appendix F: Stimuli used in Experiment 6

F	L	N	P	FL	PF	FL	FP	LN
/esy/	/eru/	/eni/	/ety/	/ebri/	/eksu/	/evry/	/efti/	/ernu/
/ɛsa/	/ɛra/	/ɛna/	/ɛta/	/ɛbra/	/ɛksa/	/ɛvra/	/ɛfta/	/ɛrna/
/osu/	/ory/	/oni/	/otu/	/obry/	/oksu/	/ovru/	/ofty/	/orni/
/ɔsã/	/ɔrã/	/ɔnã/	/ɔtɔ/	/ɔbrã/	/ɔksã/	/ɔvrã/	/ɔftã/	/ɔrnã/
/øsi/	/øru/	/øny/	/øti/	/øbru/	/øksy/	/øvri/	/øftu/	/ørny/
/œsa/	/œra/	/œna/	/œta/	/œbra/	/œksa/	/œvrɔ/	/œfta/	/œrna/
/evi/	/eny/	/emi/	/edu/	/egly/	/epsi/	/ezlu/	/ezby/	/elmi/
/ɛva/	/ɛna/	/ɛmɔ/	/ɛdaɔ/	/ɛgla/	/ɛpsã/	/ɛzlã/	/ɛzbã/	/ɛlma/
/ovi/	/onu/	/omy/	/odi/	/oglu/	/opsy/	/ozli/	/ozbu/	/olmy/
/ɔva/	/ɔnã/	/ɔma/	/ɔda/	/ɔgla/	/ɔpsa/	/ɔzla/	/ɔzba/	/ɔlma/
/øvy/	/øni/	/ømu/	/ødy/	/øgli/	/øpsu/	/øzly/	/øzbi/	/ølmu/
/œvã/	/œnã/	/œmã/	/œdã/	/œglã/	/œpsã/	/œzlã/	/œzbã/	/œlmã/