**University of Alberta**

# Lexical and Phonotactic Cues to Speech Segmentation in a Second Language

by

Manuel Sinor ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics

Edmonton, Alberta
Fall 2006

*Your file Votre référence*
*ISBN: 978-0-494-23107-4*
*Our file Notre référence*
*ISBN: 978-0-494-23107-4*

## ABSTRACT

This doctoral study explores the roles of lexical and phonotactic cues in speech segmentation (i.e., the perceptual division of speech into words) by adult Francophone learners of English.

The main, word-spotting experiment uses nonsense strings carrying real target words (e.g., 'voo*cman*'). Back-up results come from auditory and visual offline segmentation tasks as well as from a lexical decision task using isolated, spoken words.

The study shows that:

1) speech segmentation in a second language (L2) gains primarily from lexical cues pertaining to the relative usage frequency of the target words, and secondarily from phonotactic cues pertaining to the alignment of syllable and word boundaries inside the carrier strings;

2) learners of higher L2 proficiency are faster, but learners of lower proficiency also prove sensitive to phonotactic cues associated with the second language;

3) L2 speech segmentation is not facilitated by lexical cognacy between the first language and the second language, and is impeded by phonotactic patterns common to these two languages.

These findings highlight the risk that a first language can interfere with L2 speech segmentation, on the phonotactic as well as lexical level. But these findings also underscore the possibility that speech segmentation, while crucial to the acquisition of a second language, may ultimately be performed in a native-like fashion.

## ACKNOWLEDGEMENTS

I wish, first and foremost, to thank my supervisor, Dr. Gary Libben.
His expertise, creativity and ability to share knowledge have been truly inspiring.

My gratitude also goes to the members of my thesis committee:
Dr. Terry Nearey, for the exactness and generosity of his scholarship;
Dr. Johanne Paradis, for her encouraging and precise feedback throughout my PhD;
Dr. Martin Beaudoin, for his continuous support to my teaching and my research;
Dr. David Watt who, on very short notice, kindly came on board as an external reader.

Special thanks go to Dr. John Newman, who helped make my PhD a livable experience as far as the thesis defense, and to Dr. Patrick Bolger, who selflessly helped out with secondary revisions as late as after the defense.

I feel honored to have learnt from such distinguished linguists as Dr. Sally Rice, Dr. John Hogan and Dr. David Beck, and to have witnessed the unflagging patience and humor of Ms. Debra Elliott and Ms. Jana Tomasovic.

And while this thesis is dedicated to my family, I owe a lot to my friends Christian, Tracy and, last but not least, Patrick and Catherine.

Je dédie cette thèse à ma grand-mère, feue Thérèse-Isabelle Richard-Emerard.

Ma gratitude va aussi à ma mère, Mylène Sinor, mon grand-père, Denis Sinor et mon grand frère, Edouard Sinor.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# CHAPTER ONE

# INTRODUCTION

## 1.1. Overview of the Study

### 1.1.1. Background

A central aspect of language comprehension is speech segmentation, the perceptual division of speech into words. Indeed, with the exception of lexicalized collocations such as ‘quick stop’, human listeners can only comprehend the input by segmenting it into distinct words. For instance, the utterance ‘quick pairing’ can only make sense if a listener divides it into the words ‘quick’ and ‘pairing’.

Speech segmentation can be assumed to proceed through at least one of two mechanisms, described below using English examples.

One mechanism is a pop-up effect whereby listeners recognize words in the input (e.g., ‘quick’ or ‘pairing’) and thus automatically deduce the boundaries of these words. Properties of previously encountered words, such as their usage frequency, can therefore be thought to provide *lexical cues* to speech segmentation. For instance, the high-frequency word ‘quick’ seems easier to segment than the low-frequency ‘quack’.

Another but potentially complementary mechanism is the perception of speech phenomena capable of offering non word-specific, *sublexical cues* to word boundaries:
a) phonetic cues (e.g., the aspiration of word-initial stop consonants, as in ‘quick pairing’ [kwɪk#p$^{h}$ɛrɪŋ]–where the number sign symbolizes the word boundary);
b) prosodic (or, rhythmic) cues (e.g., the presence of strong syllables at the onsets of most English content words, as in ‘quick pairing’ [’kwɪk#’pɛrɪŋ]);
c) phonotactic cues (e.g., the fact that English consonants [k] and [p] may not legally cluster into a syllable onset and may therefore only cooccur in a heterosyllabic fashion, thus creating an alignment of syllable and word boundaries in ‘quick pairing’ [kwɪk.#p$^{h}$ɛrɪŋ]–where the dot sign symbolizes the syllable boundary).

### 1.1.2. Issues under investigation

The study reported in this dissertation investigates the following issues:

1) whether second language (L2) listeners use lexical and phonotactic segmentation cues;
2) which of these two cue types dominates in L2 speech segmentation;
3) whether this pattern evolves over the time course of L2 acquisition;
4) whether the use of L2 cues is influenced by equivalent cues in the first language (L1).

### 1.1.3. Rationale

#### 1.1.3.1. Rationale for the investigation of L2 speech segmentation

L2 speech segmentation is difficult, as L2 listeners, unlike L1 listeners, often hear entangled sounds in place of distinct words. Anecdotal evidence, nonetheless, suggests that segmentation improves as learners gain L2 proficiency. This study may thus help explain how a central problem of L2 performance tends to resolve itself gradually.

Actually, second language acquisition (SLA) or, more precisely, the acquisition of L2 structures (i.e., phonetic segments, morphemes, words, phrases) should logically call for the segmentation of L2 input into such structures. And yet little has been done to test the relationship between L2 acquisition and L2 segmentation (Carroll, 2002: 228). The present study attempts to address this research gap, and also seeks to find guidelines for reemphasizing segmentation skills in L2 listening training.

#### 1.1.3.2. Rationale for the investigation of L2 lexical cues

Speech segmentation can be thought to rely primarily on the word recognition, pop-up effect described above. Indeed, speech processing at large seems more driven by the search for meaningful lexical units than by the detection of incidental, sublexical cues. Also, L2 segmentation seems to get easier as L2 words become more familiar: the more often learners of English encounter the word ‘ocean’, the more easily they should detach it from ‘is’ in ‘the ocean is rising’. A lexicalist perspective, in this light, calls for an assessment of the role of lexical cues (notably word frequency) in L2 segmentation.

1.1.3.3. Rationale for the investigation of L2 sublexical-phonotactic cues

*Rationale for the investigation of L2 sublexical cues at large.* Theoretical discussions of sublexical segmentation cues have been sporadic in the SLA literature (Tarone, 1974; Hieke, 1987, 1989; Champagne-Muzar, 1991; Carroll, 2002, 2004). On an empirical level, the relevance of these cues has been overshadowed by evidence that L2 segmentation could be hindered by sublexical phenomena, such as the reduction of word-final consonants in English (as in 'wannem' [wɑ̃nə], from "I want him to go') (Henrichsen, 1984; Ito, 2001)[1] or the resyllabification of word-final consonants in French, (as in 'petit ami' [pøti.tɑmi], 'boyfriend') (Matter, 1986; Dejean de la Bâtie & Bradley, 1995). Meanwhile, studies of L2 sublexical cues have been sparse, and have left unresolved some key issues: whether phonetic cues can help L2 segmentation (Section 2.2.2.2); whether L2 listeners can use L2-specific rather than L1-based prosodic and phonotactic cues (Sections 2.2.2.3 and 2.2.2.4); how the use of L2 sublexical (notably, phonotactic) cues may evolve over the time course of L2 acquisition (Section 2.2.2.4).

*Rationale for the investigation of L2 sublexical-phonotactic cues in particular.* Weber (2000) and Weber and Cutler (2004, 2006) have pioneered this investigation, offering evidence that L2 listeners were sensitive to both L1 and L2 phonotactic patterns. But they have not tested, as this study will, the interaction of phonotactic and lexical cues in L2 segmentation, nor the weights of such cues at different levels of L2 proficiency.

Phonotactic cues, in fact, are readily testable. Unlike phonetic and prosodic factors, which are utterance-specific (and possibly intertwined: see Section 2.2.2.3), phonotactic generalizations (like the legality/admissibility vs. illegality/inadmissibility of the [bl] vs. [tl] onsets in English) apply across the board in a given language. Cluster phonotactics can therefore affect online segmentation in consistently predictable ways. For instance, the illegality of onset clusters like [kp] in English, hence the alignment of syllable and word boundaries in carrier strings like 'quick pairing' [kwɪk.#pʰɛrɪŋ], have been repeatedly shown to facilitate word-spotting (Sections 2.1.3.3 and 2.2.2.4).

[1]The English teaching literature also emphasizes these challenges (see Rosa, 2002 for review).

Phonotactic segmentation strategies also face three challenges. First, consider the example sequence 'bus ten'. Abstraction made of [t] aspiration (which could cue a syllable boundary after 'bus'), the legal onset [st] promotes the syllabification [bʌ.stɛn], hence a misalignment of syllable and word boundaries and a likely source of word-spotting difficulties (Section 2.1.3.3). Second, a CC cluster may only act as a cue if its members are perceptually distinct: so it is with [kp] in 'quick pairing', but not with the homorganic [kg] in 'mock guide'. Third, phonotactic cues may be weakened by L1-L2 phonotactic differences, as with [pn], which is rare in French (e.g., 'pneu' [pnø], 'tire') but fully illegal in English. Ensuing cases of missegmentation are easy to foresee (e.g., '*be pnoton' for 'beep not on'), especially since L2 listeners often add epenthetic vowels to L2 clusters that violate L1 phonotactics (Dupoux, Kakehi, Hirose, Pallier & Mehler, 1999; Kabak, 2003). Altogether, these limitations of phonotactic cues highlight the possibility, crucial to this research, that L2 listeners may rely on the integration of additional, lexical segmentation cues.

### 1.1.4. Overall methodology

This study tested segmentation skills among learners of English as a Second Language (ESL), who had a Francophone or Chinese background and were based in Edmonton or Montréal. A main, word-spotting task used word-ending nonsense strings like 'lous<u>ripe</u>'. Back-up results came from auditory and visual offline segmentation tasks and a lexical decision task using isolated, spoken words.

### 1.1.5. Independent variables

#### 1.1.5.1. Subject's L2 Proficiency

L2 proficiency was primarily understood as proficiency in L2 word recognition, and was thus operationalized as accuracy scores in lexical decision and word-spotting.

Two proficiency levels were considered: Intermediate vs. Advanced.

1.1.5.2. Subject's L1

Francophone learners were the main population of interest, for two reasons. Lexically, English and French can be similar (e.g., 'post'-'poste') or different (e.g., 'ripe'-'mûr'). Phonotactically, similarities (such as the legality of onset [bl]) also coexist with differences (such as the fact that onset [sn] is common in English but not in French – see corpus-based analysis in Section 3.1). English and French therefore offered a fertile ground for testing how L1 lexicon and L1 phonotactics might affect L2 segmentation.

English and Mandarin are dramatically more different. Mandarin has borrowed few words from English (e.g., 'card', translated in Mandarin as 'ka')[2], and while sharing some consonant-glide onset clusters (e.g., [kw, pj]) with English, it lacks the obstruent-obstruent and obstruent-liquid onsets of English (e.g., [sn, pl]). A Chinese control group seemed therefore ideal for testing L1 French transfer on L2 English segmentation.

1.1.5.3. Frequency

How familiar one is with a word normally reflects the extent to which one uses it. Therefore, in principle, L2 word familiarity could have been measured by using learners' ratings of how frequently they used the relevant words in their L2. L1 frequency ratings were used instead, for three reasons: 1) the lack of accessible corpora with L2 frequency ratings; 2) the statistical noise that would have derived from using frequency ratings by all participants; 3) the likely proportionality of L2 and L1 frequency data, as words of high L1 frequency are likely to be taught early during programs of L2 teaching.

The Frequency variable gave rise to two types of English words: 1) high-frequency (HF) words, like 'cash'; 2) low-frequency (LF) words, like 'cure'.

1.1.5.4. Cognacy

Words with French cognates (or YESCOG), like 'list' (in French, 'liste' [list]) were opposed to words without (or NOCOG), like 'mad' (in French, 'fou' [fu]).

---

[2]Hall-Lew (2002) recensed no more than 112 English loanwords in a study of monolingual Mandarin speakers from the provinces of Yunnan, Beijing, Xi'an and Taiwan.

1.1.5.5. Onset Type

The variable Onset Type was operationalized via the notion of onset markedness. Onset markedness was understood as the relative frequency by which a given CC onset cluster occurred within preselected corpora of English and French monosyllabic words. Thus, as explained in a pilot corpus analysis (Section 3.1), a given CC onset would be labelled 'marked' if its occurrence frequency value fell in the 0.3% - 3% range, and would be labelled 'unmarked' if its frequency value was above 3%.

On this basis, the English words and English-like nonwords used for the study displayed four types of onsets:

1) C onsets, that is, simple onsets (as in 'kite' or 'riss');

2) CC1 onsets, which are unmarked in English and French (as in 'black' or 'proun');

3) CC2 onsets, which are marked in French but unmarked in English ('skin' or 'spile');

4) CC3 onsets, which are illegal in English and French (as in 'tmoul').

### 1.1.6. Research hypotheses

These hypotheses were treated as two-tailed because of the exploratory nature of the study. They were also grounded in the literature on L2 speech segmentation, which explains why their justifications punctuate the review of this literature (Section 2.2).

*Hypothesis 1*

*L2 lexical factors (pertaining to the Frequency variable) and L2 phonotactic factors (pertaining to the Onset Type variable) affect L2 speech segmentation.*

*Hypothesis 2*

*L1 lexical factors (pertaining to the Cognacy variable) and L1 phonotactic factors (pertaining to the Onset Type variable) affect L2 speech segmentation.*

*Hypothesis 3*

*L2 proficiency factors (pertaining to the Proficiency variable) affect L2 speech segmentation, along one of two alternative developmental paths:*

- *Subhypothesis 3a. Learners gradually shift from phonotactic to lexical cues.* Here is the rationale for this subhypothesis. Phonotactics seems a tighter hence faster-to-grasp domain than vocabulary[3], and so less proficient learners are more likely to be phonotactically than lexically skilled. They may also be prompt to use phonotactic cues, so as to make up for lexical deficiencies. However, Advanced learners, having wider L2 vocabulary, may have become so sensitive to lexical cues as to give up phonotactic cues.

- *Subhypothesis 3b (alternative). Learners mix lexical and phonotactic cues increasingly.* Here is the rationale for this subhypothesis. Advanced learners, having been more exposed to L2 lexical and phonotactic cues, may be more likely than Intermediate learners to integrate such cues. This strategy, rather than making up for L2 deficiencies, may fill a need to accumulate opportunities of maximal efficiency in speech processing.

### 1.1.7. Aim of the study

Given these hypotheses, this study assesses the contributions of L1 and L2 lexical and phonotactic cues to the performance and the acquisition of L2 speech segmentation.

## 1.2. Structure of the Dissertation and Investigation

*Chapter 2.* This chapter reviews the literature relevant to this research.

Section 2.1 focuses on monolingual (L1) speech segmentation, for three reasons: 1) the above hypotheses, while grounded in the L2 literature, were also inspired by online and developmental L1 patterns; 2) L1 skills should be the ultimate target for L2 listeners; 3) speech segmentation has received more attention in the L1 than L2 literature.

In Section 2.2, past studies on lexical and sublexical cues to L2 segmentation evoke the possibility that L2 learners may also favor cue integration increasingly.

*Chapter 3.* This chapter reports a pilot test of phonotactic cues in L1 English.

[3]This assumption implies that phonotactic generalizations (about the admissibility or non-admissibility of sound sequences in a given language) can refer to the syllabic structures rather than to some word-level exemplars of the relevant sound sequences (see Section 2.1.3.3).

Section 3.1 presents the rationale for this pilot test.

Section 3.2 reports a statistical analysis of CC clusters in English and French. This analysis served to: 1) highlight the correlation between the phonotactic status and the occurrence frequency of individual clusters in these two languages; 2) identify clusters to be tested in the following pilot experiments.

Sections 3.3 and 3.4 report Experiments 1 and 2. These offline experiments, as summarized below, converged to attest the potential importance of phonotactic information in English speech segmentation. They also helped to identify onset CC clusters that might best elicit such phonotactic effects during subsequent, main experiments with L2 learners of English.

Experiment 1 tested whether native Anglophones, when asked to segment nonsense CVCCCVC strings like 'nantlis' [nantlɪs], relied on the phonotactic status of string-medial CC clusters (e.g., legal offset [nt] vs. illegal onset/offset [tl]). Printed strings served to prevent mishearings of the string-medial CC clusters and thus yield transparent insights into the phonotactic acceptability of these clusters in English.

Experiment 2 reused these stimulus strings in the auditory mode to test the actual contribution of the relevant CC clusters to English speech segmentation.

*Chapter 4.* This chapter reports the main experiments of the study, as they pertain to speech segmentation in L2 English.

Section 4.1 discusses, in greater depth than would be possible within the constraints of the subsequent experimental sections, the participants and procedures involved in these main experiments.

Sections 4.2 to 4.5 report Experiments 3 to 6, as summarized below.

Experiment 3, a lexical decision task, tested how ESL learners processed isolated English CVC(C) words (e.g., 'page') and English-like CCVC nonwords (e.g., 'broof'). The aim was to preassess, on an item-by-item basis, the participants' sensitivity to the lexical cues (i.e., Frequency and Cognacy) and phonotactic cues (i.e., Onset Type) to be tested in the main, online segmentation task (Experiment 4).

Experiment 4, a word-spotting task, tested how L2 learners processed the target words of Experiment 3 (i.e., words of high vs. low frequency and cognacy vs. no-cognacy with French) but now in the context of segmentable CVCCVC(C) strings (e.g., 'coafpage' vs. 'koespage') built on the CC onsets of contrastive phonotactic statuses displayed by the nonwords of Experiment 3 (e.g., legal [br] vs. illegal [sr]).

Experiment 5, an offline word-likeness judgment task, served as control for Experiment 4. It asked participants to decide how acceptable each of the CVCCVC(C) strings of Experiment 4 would be as novel English words. The aim was to check the extent to which Experiment 4, an experiment originally designed to tap the retrieval of phonological lexical entries, might have been inadvertently affected by semantic noise resulting from the presence of real words within the word-spotting stimuli.

Experiment 6, an offline segmentation task, also served as control for Experiment 4. It sought to double-check the phonotactic biases that had emerged from the word-spotting performances of Experiment 4. It used CVCCCVC(C) strings built on the CC onset clusters previously tested through the word-spotting technique. It also used the same visual task of off-line segmentation as in Experiment 1, the visual format allowing to prevent mishearings of the CC clusters of interest.

*Chapter 5.* This last chapter recapitulates the study, with a focus on the two main conclusions of the main online experiments (Experiments 3 and 4), namely:

1) L2 phonotactics helps; L2 word frequency helps more;

2) L1 cognacy and L1 phonotactics do not help.

The chapter, finally, explores the potential implications of these conclusions for an understanding of how second languages are processed, learnt and taught.

# CHAPTER TWO

# LITERATURE REVIEW

## 2.1. LEXICAL AND SUBLEXICAL FACTORS IN MONOLINGUAL SPEECH PROCESSING

### 2.1.1. Lexical factors

2.1.1.1. The primary importance of words in monolingual speech processing

Psycholinguistic research consistently shows that words are fundamental units of L1 speech processing, at every stage of language development. Infants previously familiarized with isolated words listen longer to sequences containing these words, and infants previously familiarized with word sequences listen longer to the single words contained in these sequences (Jusczyk & Aslin, 1995). Adults continue to display strong lexical sensitivity in online speech processing, as shown by their tendency to respond to words faster than to nonwords during lexical decision tasks (e.g., Rubenstein, Garfield, & Milikan, 1970) and by their ability to spot real words inside nonsense materials during word-spotting tasks (e.g., Bacri & Banel, 1995; McQueen, 1998).

2.1.1.2. Lexical factors of influence on spoken word recognition

*Influential properties of target words.* “Target” means here “word to be recognized”.

Target word properties that can influence spoken word recognition include: word length (e.g., Baddeley, Thomson & Buchanan, 1975), uniqueness point[4] (e.g., Radeau, Morais, Mousty & Bertelson, 1989), word concreteness (i.e., the relative concreteness of a word’s meaning of a word–see, e.g., Kroll & Merves, 1986), word imageability (i.e., the degree to which a word arouses a mental image–see Tyler, Voice & Moss, 2000).

[4]The uniqueness point is where a word becomes uniquely distinguished from similar words. Thus, the uniqueness point for ‘elephant’ is [f], first sound to distinguish ‘elephant’ from ‘elegant’ or ‘element’.

Word usage frequency is also influential. For isolated words, frequency increases trigger increases in lexical decision accuracy and speed (e.g., Marslen-Wilson, 1990; Taft & Hambly, 1986), word naming speed (Balota & Chumbley, 1985), and eye fixation time during eye-tracking experiments that associate stimulus pictures to the target words (Dahan, Magnuson & Tanenhaus, 2001). Similarly, for words embedded in larger stimuli (typically, nonsense syllable sequences/strings), frequency effects trigger faster word processing (Radeau, Morais & Ségui, 1995; Liu, Bates, Powell & Wulfeck, 1997). These frequency effects are generally thought to reflect effects of word familiarity (i.e., how familiar a word is to a listener). Thus, while admitting that familiarity ratings can best reflect a speaker's perceived experience of specific words, researchers typically view frequency as a more readily available statistic that correlates highly with rated familiarity (see, e.g., Kreuz, 1987; Kacinik, Shears & Chiarello, 2000; Sobkowiak, unpublished).

*Influential properties of words present in the environment of a target word.* Such words can facilitate the recognition of a target (e.g., 'blues') if they relate semantically or phonologically to it (e.g., 'jazz' or 'blows'– see e.g., Slowiaczek & Hamburger, 1992).

*Influential properties of words absent from the environment of a target word.* Such words, if they relate to the target (e.g., phonologically or orthographically) can vie with it for recognition. This phenomenon (aka. lexical competition in word recognition) is best evidenced via a lexical variable known as neighborhood density (i.e., the amount of words that resemble a target word). Consider, thus, the words 'cat' and 'boil'. 'Cat', with similar-sounding words like 'fat', 'rat', mat', 'sat', 'cut', 'kit', has a dense neighborhood, while 'boil', with fewer similar-sounding words, like 'box', 'coil' or 'soil', has a sparse neighborhood. Evidence of slower and less accurate responses to dense- than sparse-neighborhood words, suggests that, indeed, word recognition involves the coactivation and competition of several word-forms. Such evidence has been found in connection with targets played in isolation during lexical decision tasks (e.g., Luce & Pisoni, 1998; Vitevitch & Luce, 1998) as well as with targets embedded in larger, word-spotting stimuli (McQueen, Norris & Cutler, 1994; Norris, McQueen & Cutler, 1995).

2.1.1.3. Summary and implications

The above discussion, by confirming words as fundamental units of speech processing, strengthens the view that word recognition (or, more precisely, lexical access) may be the force driving speech segmentation. The above findings have also highlighted variables susceptible of influencing the recognition/segmentation of string-embedded words. Frequency would become an independent variable in the present study. And while other variables, like word concreteness or imageability, could not possibly all be balanced across the stimuli, efforts were made to control two variables likely to affect lexical decision and word-spotting: phonological and orthographic neighborhoods.

**2.1.2. From a lexical approach to a multiple-cue approach to speech segmentation**

Speech segmentation, as discussed earlier, can in principle be viewed as a mere by-product of lexical access, with no need for cues other than the properties of the target words. This view is all the more legitimate since acoustic markers of word boundaries (i.e., vocal fry, glottalization, laryngealization; see Quené, 1992 for review) are small and inconsistent (Lehiste, 1972; Nakatani & Dukes, 1977). Indeed, the anecdotal experience of speech as a continuous phenomenon suggests that the sublexical cues to be described next (i.e., phonetic, prosodic and phonotactic cues) are at least subtle, at worst unreliable.

Sublexical phenomena, nonetheless, may together make up for their respective shortcomings, while at the same time facilitating speech segmentation. This assumption gains strength if one holds that speech segmentation and, more generally, speech processing, are guided by maximal efficiency. Listeners engaged in casual conversation can indeed detect words spoken at a rate of no less than 20 to 30 phonemes per second (Cole & Jakimik, 1980). In terms of efficiency, thus, it is hard to imagine that a communicative task as vital as the segmentation of speech into words would not exploit cues directly relevant to its operation. There is, therefore, merit to the view that speech segmentation, while primarily driven by lexical access, may exploit an integrated array of additional, bottom-up/sublexical cues. This assumption, which can be labelled 'multiple-cue integration in speech segmentation', retains its appeal whether one considers lexical access under a sequential approach or a competition-based approach.

Under a sequential approach, as in the Cohort Model (e.g. Marslen-Wilson & Welsh, 1978; Marslen-Wilson, 1990), lexical access and its speech segmentation corollary proceed sequentially, as follows: 1) an incoming utterance (e.g., 'an elephant came' [ənɛlɛfəntkejm]) is processed by accessing all possible words starting with adjacent input segments (e.g., 'elegant' [ɛlɛgənt], 'element' [ɛlɛmənt], 'elephant' [ɛlɛfənt]); 2) a target word (e.g, 'elephant' [ɛlɛfənt]) is accessed/recognized when its uniqueness point is reached (i.e., [f]); 3) recognition of this word makes evident the end of this word and the onset of the next in the utterance (i.e., 'came' [kejm], here). The appeal of this scenario is that it fits with the experience of hearing words one by one. And yet one may wonder how the same procedure would recover from a momentary failure if no additional word boundary cue was available from which to restart the whole sequential process. Suppose that, in 'an elephant came', 'elephant' went unrecognized. In the present scenario, the onset of 'came' would then go undetected and segmentation would fail. But if, efficiency oblige, the processor had access to extra, sublexical cues such as the word-initial aspiration of [k] in 'came' [k$^h$ejm], then speech segmentation could operate smoothly (cf. Frauenfelder, 1985 for a similar argument).

Under a competition-based approach, the same reasoning applies. Here again, the assumption is that word boundaries emerge–and thus segmentation becomes possible–as a by-product of lexical access. But unlike sequential models, lexical competition models like TRACE (McClelland & Elman, 1986) hold that competition occurs anywhere in the input rather than sequentially. For instance, in the context of the input [kætəlog], the target 'catalogue' would compete with words placed at the start the input (like 'cat') and with words placed at the end of the input (like 'log'). The appeal of this scenario is that the neighborhood effects reviewed earlier do indeed suggest the ubiquity of lexical competition during lexical access. However, the notion of maximal efficiency also suggests that a segmentation model based on lexical competition anywhere in the input would gain from extra word boundary cues. This suggestion came from the authors of TRACE themselves (ibid., 1986: 63-64). They argued that, for the target 'party' to lose out against 'par tea' (and, one may add, notwithstanding additional top-down/semantic cues), there should be a bottom-up cue like a pause between 'par' and 'tea'.

Overall, then, lexical access and its speech segmentation corollary seem to find maximal efficiency in a synergy of lexical competition mechanisms and perceptible sublexical word boundary phenomena. Put differently, speech segmentation may in principle be assumed to involve the integration of lexical and multiple sublexical cues.

This assumption has already been computationally, experimentally and theoretically supported. On a computational level, simulations have shown that speech segmentation procedures could be guided by both lexical and sublexical cues (e.g., Grossberg & Myers, 2000; Norris, McQueen, Cutler, & Butterfield, 1997). On an experimental level, there has been evidence that English speech segmentation depended both on lexical competition and on a metrically-guided procedure baptized Metrical Segmentation Strategy (MSS), whereby strong syllables are postulated to be the onsets of words (Cutler & Norris, 1988; McQueen, Norris & Cutler, 1994). On a theoretical level, the Shortlist model (Norris, 1994) holds that this empirically attested MSS procedure actually guides lexical access (hence, speech segmentation) by boosting the activation of word recognition candidates that start with strong syllables. Similarly, the Good Start model posits that while the discovery of word boundaries is primarily lexically driven, sublexical cues such as allophonic variation or metrical stress can get lexical activation off to a "good start" by making some word onsets more salient (Gow & Gordon, 1995).

Against this context, the next section reviews the sublexical cues attested so far.

### 2.1.3. Sublexical cues

#### 2.1.3.1. Sublexical-phonetic cues

Phonetic segmentation cues are allophonic variation and segmental duration.

Allophonic variation can facilitate speech segmentation in English. For instance, aspiration of syllable-initial stops can help listeners distinguish 'gray twine' [grej#$t^h$wajn] from 'great wine' [grejt#wajn] (Nakatani, unpublished, cited in Church, 1987), while variation between the syllable-initial [l] (describable as light and glide-like) and the syllable-final [l] (describable as dark and syllabic) can help listeners distinguish between 'we loan' and 'we'll own' (Nakatani & Dukes, 1977).

Durational cues typically involve the fact that consonants tend to be longer word-initially than word-medially (e.g., Christie, 1977; Eefting, 1991; Quené, 1992). Christie found that English listeners perceived [hɛlpəsnejl] with a word-initial [n] (as 'help us nail') when the [n] was lengthened, but perceived the same sequence with a word-medial [n] (as 'help a snail') when the [n] was shortened. Yet such cues are not without flaws. Sinor (2003) found that, even after durational contrasts between short/resyllabified vs. long/word-initial versions of French [t] had been artificially enhanced, L1 Francophones had trouble differentiating between phrases like 'petite amie' ([pøti.t#ami], 'girlfriend') vs. 'petit tamis' ('little sieve', [pøti#.tami])[5].

In short, allophonic and durational cues may not suffice to resolve the speech segmentation problem (see Jusczyk & Luce, 2002: 23 for a similar argument).

### 2.1.3.2. Sublexical-prosodic cues

Prosodic cues derive from the regularity of rhythmic patterns in some languages. Thus, in English, primary stress affects most polysyllabic content words (e.g., *mo*torbike or *ta*ble – cf. also Cutler & Carter, 1987) and, ultimately, speech segmentation (Cutler & Norris, 1988; Cutler & Butterfield, 1992; Norris, McQueen & Cutler, 1995). Cutler and Norris thus found that participants spotted 'mint' more slowly in doubly stressed bisyllables (e.g., 'mínteív') than in bisyllables with primary stress only (e.g., 'míntef'). On this basis, they posited a Metrical Segmentation Strategy (MSS) whereby English listeners treated all stressed syllables as possible word onsets. The MSS can explain how listeners, by segmenting 'minteiv' into the stressed/word-like 'min' and 'teiv', slowed down their detection of the single target 'mint'.

This finding, however, could have reflected a phonetic rather than prosodic cue (i.e., the aspiration of syllable-initial voiceless stops, as in 'minteif' [mɪn.tʰejf]), as pointed out by Davis (2000: 108-109). Prosodic and phonetic cues thus seem to pose two research problems: they are difficult to disentangle, and they closely depend on the utterance in which they occur. Phonotactic cues, meanwhile, seem easier to test.

---

[5] Durational cues did help segmentation when [t] was enhanced to simulate a geminate consonant (as in 'petite Tammy' [pøtittami], 'little Tammy'), thus contrasting with the singleton [t] of the above examples.

### 2.1.3.3. Sublexical-phonotactic cues

*Preamble: categorical vs. probabilistic phonotactics.* Phonotactics refers to language-specific restrictions on the sequencing of phonetic segments within and between words.

Categorical phonotactics depicts segment arrangements as legal or illegal (cf., e.g., Clements & Keyser, 1983). Thus, English words may start in str- (e.g., 'string') but not in stl-. This categorization can be restated in syllabic terms, namely the legality vs. illegality of str- vs. stl- as syllable onsets in English[6]. Such depictions not only reflect the way that monolinguals segment speech before the age of 10 months (Jusczyk, Luce & Charles-Luce, 1994), they also help design contrastive levels of difficulty in the context of segmentation tasks. For instance, as further explained at the end of this section, the segmentation of the nonsense string 'zun<u>car</u>' should gain from the alignment of word and syllable boundaries caused by the *illegal* onset [nk], while the segmentation of 'zus<u>car</u>' should suffer from the misalignment of boundaries caused by the *legal* onset [sk].

Probabilistic phonotactics addresses the difficulty of categorizing certain segment sequences as strictly legal vs. illegal. For instance, the French onset [pn] would seem fully illegal were it not for its rare occurrence in words like 'pneu' [pnø] ('tyre'). In English, also, graded rather than absolute phonotactic distinctions are worthwhile. One may thus talk of high-probability patterns (like the occurrence of [st] as onset, in 'star', stitch', 'stop', 'stung') and low-probability patterns (like the occurrence of [sf] as onset, in 'sphere' or 'Sphinx'). Probabilistic phonotactics, in this light, has become the favored approach to phonotactics in psycholinguistic research (see, e.g., Kessler & Treiman, 1997; Vitevitch & Luce, 1999; Vitevitch, Luce, Charles-Luce & Kemmerer, 1996).

*Insights into the phonotactics of word processing.* This study rests on the assumption that phonotactic and lexical effects on word processing are distinct (yet possibly interactive).

To assess the validity of this distinction, one may wonder whether phonotactic, probability effects on word processing can be dissociated from lexical, neighborhood

[6]See, however, Blevins (2003) for a challenge to the widely held assumption that phonotactic constraints are based on syllable structure and for a suggestion that consonant sequencing is instead conditioned by phonetic factors such as manner of consonant articulation.

density effects. At first glance, this may seem unlikely since the higher the phonotactic probability, the denser the neighborhood density (Vitevitch & Luce, 1998: 325). And yet Luce and Large (2001) did find that probability and density effects could be disentangled. Their same-different match task[7], while producing faster responses to high- than low-probability words, produced slower responses to high- than low-density words, and no significant interaction between probability and density.

A next issue of interest, then, is whether phonotactic probability vs. neighborhood density affect sublexical vs. lexical levels of speech processing. Such insights can be gained from extending the present discussion to the issue of nonword processing, given that the sublexical level of representation is common to words and nonwords.

*Insights into the phonotactics of nonword processing.* The focus here is on the 1999 study of a round of investigations by Vitevitch and Luce (1997, 1998, 1999, 2005).

A same-different matching task was first run, which covaried the phonotactic probability and neighborhood density of word and nonword stimuli, so as to reflect the natural positive correlation between these two variables. Nonwords were mixed with words in order to foster sublexical processing across the board. Responses to nonwords were significantly faster for high density-probability nonwords than for low density-probability nonwords. This result was taken as evidence that nonword processing had not been hindered by density increases and had been facilitated by probability increases.

A lexical decision task was then run. Lexical decision forces the assessment of nonwords against words, hence a lexical rather than sublexical way to process nonwords. The reverse of the above result was therefore expected, namely slower responses for high than for low density-probability nonwords. This did happen, confirming that when probability increases had helped nonword processing, they had done so sublexically.

In sum, this study (and its 2005 replica, using new, duration-matched stimuli), supports the scenario put forward by Vitevitch and Luce (1997, 1998) whereby nonwords are primarily processed at the sublexical level, the level at which phonotactics

[7]Such a task requires, on a given trial, to determine as fast and as accurately as possible if two stimuli are the same or different. The dependent variable, here, was the speed at which participants responded 'same'.

operates. This conclusion motivates the decision, in the present study, to examine phonotactic effects on nonword processing during a lexical decision task (Experiment 3).

The above discussion also strengthens the motivation of the present research to compare lexical vs. sublexical-phonotactic cues in L2 speech segmentation. Indeed, given the above indication that phonotactics affects nonword processing sublexically, given also the possibility to describe phonotactic patterns using syllabic and articulatory rather than lexical criteria, and given that models of word recognition (e.g., Shortlist, TRACE) tend to separate lexical vs. sublexical processing levels, one may confidently restate the conceptual distinction between lexical and sublexical-phonotactic factors in word recognition. If, finally, one reiterates that word recognition drives speech segmentation, then the above literature appears to legitimize the present investigation of "lexical vs. phonotactic cues to speech segmentation".

*Phonotactic cues to the auditory segmentation of string-embedded words.* The following discussion is couched in the terminology of categorical rather than probabilistic phonotactics, because the categorical standpoint has been more common in word-spotting studies and thus seems more suitable to explain the ins and outs of such studies.

A study by Norris, McQueen, Cutler and Butterfield (1997) highlights the background for later studies of phonotactic effects on the L1 segmentation of string-embedded words (e.g., McQueen, 1998; Mattys, Jusczyk, Luce, & Morgan, 1999; Mattys & Jusczyk, 2001; Dumay, Frauenfelder & Content, 2000, 2002). On a methodological level, this study shows that, by having to extract real words from nonsense contexts (e.g., 'apple' in 'fapple' or 'vuffapple'), listeners are obliged to perform a task vital to the skill of speech segmentation, namely the identification of the boundaries of a target word. On an empirical level, the possibility for phonotactics to affect speech segmentation emerges from the finding that English listeners spotted 'apple' faster and more accurately in 'vuffapple' (where the leftover syllable 'vuff' is phonotactically legal) than in 'fapple' (where the leftover 'f' is not a legal syllable). On a theoretical level, such a finding can justify a formal account of phonotactic segmentation effects such as Norris et al.'s proposal of a Possible Word Constraint (PWC). The PWC operates by inhibiting the recognition of a word (e.g., 'apple', activated from the input 'fapple') if this recognition wrongfully implies that the rest of the relevant input is a phonotactically viable word (e.g., 'f'). As Norris et al. pointed out, this mechanism is compatible with a lexical, competition-based approach to speech segmentation. Indeed, the PWC constraints the competition of candidates for word recognition (e.g., 'apple', 'fapple', 'f'). And computer simulations of the PWC within the framework of the Shortlist model (Norris, 1994) showed that the PWC and lexical competition mechanisms could together account for experimental data from Norris et al.'s study and past studies of speech segmentation. Here then, in line with Section 2.1.2, is an indication that lexical and sublexical-phonotactic considerations fit together within an integrated, multiple-cue approach to speech segmentation.

McQueen (1998), while confirming the effects of phonotactics on speech segmentation, introduced in the word-spotting paradigm the 'Aligned' and 'Misaligned' conditions, which are central to the word-spotting task of the present study. Prior to testing the role of phonotactic cues in Dutch, McQueen reasoned that two-segment sequences (or biphones) were bound to signal syllable boundaries if they could not legally appear inside Dutch syllables. Thus, [mr] is not a legal syllable onset or offset cluster in Dutch. It must therefore be syllabified into two heterosyllabic consonants and thus be split by a syllable boundary (symbolized by a dot) in a nonsense string like [fim.rɑk] (where 'rok' [rɑk] means 'skirt'). In contrast, [dr] is a legal syllable onset in Dutch. It must therefore appear in onset position, after the syllable boundary, in a nonsense string like [fi.drɑk]. These background observations gave rise to the Aligned and Misaligned conditions, as reexplained below using English examples.

The 'Aligned' condition involves target-bearing strings like 'zunlaugh', built on phonotactically illegal CC clusters (e.g., [nl], illegal onset or offset in English). An illegal CC cluster like [nl] is split by a (phonotactic) syllable boundary, as shown by the transcription [zʌn.læf]. It turns out, in the same example, that this syllable boundary aligns with a lexical boundary, namely the onset of the target word 'laugh'. This alignment is indicated by the juxtaposed dot (syllable boundary) and number sign (word boundary) in [zʌn.#læf].

The 'Misaligned' condition involves target-bearing strings like 'zoklaugh', built on phonotactically legal CC clusters (e.g., [kl], legal onset in English). A legal CC onset like [kl] must follow a syllable boundary, as shown by the transcription [zɔ.klæf]. Here, the syllable boundary is misaligned with a lexical boundary (i.e., the onset of 'laugh'). This misalignment is shown by the mismatch of the dot and number sign in [zɔ.k#læf].

McQueen (1998) found that Dutch target words were easier to spot in the Aligned than Misaligned condition. This finding was later confirmed for French (Crouzet & Bacri, 1998, 1999a; Dumay, Content & Frauenfelder, 1999; Dumay, Frauenfelder & Content, 2000, 2002) and English (Kirk, 2000; Weber, 2000). Thus, English 'laugh' should be easier to spot in 'zun.laugh' than in 'zo.klaugh'. In phonotactic terms, this pattern amounts to the apparent paradox that the word segmentation of a target-bearing string is facilitated if the string-medial cluster is illegal, but inhibited if the cluster is legal. Table 1 restates this conclusion, along with the trademarks of the Aligned vs. Misaligned word-spotting conditions.

Table 1. Summary of the functioning of phonotactic segmentation cues

| Phonotactic stimulus condition | Example of word-final string | String-medial CC cluster of interest | Cluster's phonotactic status | Availability of phonotactic segmentation cues (i.e., facilitatory segmenting effects) |
|---|---|---|---|---|
| Aligned | zuk<u>pit</u> | [kp] onset | illegal | Yes |
| Misaligned | zu<u>flame</u> | [fl] onset | legal | No |

Two observations can specify this contrast between the facilitatory vs. inhibitory effects of the Aligned vs. Misaligned conditions.

First, the segmentation benefits of illegal over legal clusters also apply to stimulus pairs like 'tee<u>nv</u>ok'-'tee<u>nt</u>ok', where targets (like 'teen') are string-initial rather than string-final, and where the relevant CC clusters vie for phonotactic legality in offset rather than onset position (i.e., illegal offset [nv] vs. legal offset [nt]). McQueen (1998) found that, in settings such as these, the Aligned vs. Misaligned conditions continued to generate facilitatory vs. inhibitory effects.

Second, there is evidence that the phonotactic status of the onset is more crucial to speech segmentation than the status of the offset. Preferences for syllables with onsets have indeed been attested typologically (Blevins, 1995) as well as perceptually (Flemming, 1995). In addition, word-spotting performance suffers more when the onset of a string-final word is misaligned with a syllable onset (e.g., Dutch 'rok', skirt, in 'fi.drok') than when the offset of a string-initial word is misaligned with a syllable offset (e.g., Dutch 'vel', skin, in 'velm.brul'). This 'onset-offset assymetry' applies to French, Dutch and English (Banel & Bacri, 1997; McQueen, 1998; Dumay et al., 2002; Gaygen & Luce, 2002). An implication of this effect is that the misalignment of word and syllable boundaries is more likely to take its toll on word-spotting if the target words are string-final. Hence, string-final targets should maximize the chances of eliciting phonotactic effects in the word-spotting task of the present study (Experiment 4).

*Summary.* Here are the main points from this discussion of phonotactics. First, phonotactics can be considered from two complementary approaches. The categorical approach, based on the legal-illegal dichotomy, is both psychologically viable (as shown with regard to infant speech segmentation) and methodologically useful (for building contrastive Aligned and Misaligned word-spotting items). The probabilistic approach, based on distributions of clusters in specific syllabic positions, is also psychologically viable (as shown by evidence of probabilistic effects in Vitevitch and Luce's studies), and conducive to fine-grained analyses of how listeners use phonotactic cues. Second, phonotactic cues and lexical cues, though they can be conceived as distinct from one another, can together affect speech segmentation, as has been shown experimentally (Luce & Large, 2001) and theoretically (Norris et al., 1997). Third, the above literature highlights two standard L1 patterns, which may or may not continue to apply to L2 listeners: 1) in lexical decision, phonotactic effects should be more significant for nonwords than for words; 2) in word-spotting, target words should be more easily spotted in Aligned than Misaligned carrier strings.

### 2.1.4. Conclusion

As way of conclusion, here is some evidence for multiple cue integration in L1 speech segmentation.

Monolingual infants can not only use phonetic cues (Hohne & Jusczyk, 1994), prosodic cues (Jusczyk, Cutler & Redanz, 1993) and phonotactic cues (Jusczyk, Luce & Charles-Luce, 1994), they can also integrate prosodic and phonotactic cues (Mattys, Jusczyk, Luce & Morgan, 1999). They have indeed two good reasons to integrate cues. First, they face a bootstrapping problem, whereby they must learn vocabulary but know no words to begin with and so need to use all the sublexical cues that they can find. Also, they face an indeterminacy problem, whereby none of the three types of sublexical cues can be held to be always available, detectable and conducive to accurate segmentation.

Monolingual adults, despite wider access to lexical cues, continue to integrate cues, within the sublexical domain and between the sublexical and the lexical domains. There is indeed evidence, from word and phoneme monitoring studies with adults, that prosodic and lexical activation effects interact (McQueen, Norris & Cutler, 1994; Norris, McQueen & Cutler, 1995; Sanders & Neville, 2000). Recall also the growing consensus that, regardless of age, sublexical cues can boost lexical access (Section 2.1.2).

In sum, speech segmentation poses monolinguals of all ages with the same problem, which likely calls for the same solution. The problem is that connected speech does not reliably mark word boundaries. The likely solution is multiple cue integration, which monolinguals seem to implement differently depending on their current linguistic abilities. Monolingual infants, lacking lexical knowledge, limit cue integration to the sublexical domain. Monolingual adults, having high lexical and sublexical/phonological knowledge, seem to apply cue integration to both the lexical and sublexical domains.

This scenario raises the issue of whether listeners who have limited lexical and sublexical knowledge also attempt to integrate lexical and sublexical cues, or whether they predominantly use lexical cues since these are more reliable than sublexical cues. Adult L2 learners are an ideal population for investigating this issue because their acquisition of both the lexicon and phonology of the target language is yet incomplete.

## 2.2. Lexical and Sublexical Factors in L2 Speech Processing

### 2.2.1. Lexical factors

#### 2.2.1.1. The primary importance of words in L2 speech processing

L2 listeners have more trouble than L1 listeners in recognizing isolated words (Bradlow & Pisoni, 1999) and sentence-embedded words (Bradlow & Bent, 2002; Mayo, Florentine & Buus, 1997; Meador, Flege & MacKay, 2000; van Wijngaarden, 2001; van Wijngaarden, Steeneken & Hougast, 2002). Yet they also show a widespread tendency to try and decode messages word by word, even though this strategy does not guarantee input comprehension (Conrad, 1985; O'Malley, Chamot & Kupper, 1989; Rost, 1990).

At the same time, the L2 teaching literature contains suggestions that the teaching of listening skills should be word-driven (see, e.g, Moudraia, 2001 for a review). It has even been argued that the acquisition of L2 words might predispose, rather than result from the development of L2 phonemic abilities (Bradlow & Pisoni, 1999).

#### 2.2.1.2. Learner-related factors of influence on L2 spoken word recognition

Age of acquisition can affect L2 spoken word recognition, with early learners recognizing L2 words faster than late learners (Izura & Ellis, 2002; Sebastian-Galles, Echeverria & Bosch, 2005). Gains of L2 proficiency can also help the recognition of isolated L2 words (Krause, 1988; Imai, Walley & Flege, 2005) as well as sentence comprehension (van Wijngaarden, Steeneken & Hougast, 2002). Hence the hypothesis, in this study, that L2 proficiency may affect L2 speech segmentation (Hypothesis 3).

#### 2.2.1.3. Word familiarity/frequency effects on L2 spoken word recognition

*Familiarity/frequency effects on the recognition of isolated L2 words.* The possibility for such effects only indirectly emerges from lexical decision evidence that L2 learners process real words faster than nonwords unheard before (Pál, 2000; Sinai & Pratt, 2002).

*Familiarity/frequency effects on the recognition/segmentation of string-embedded L2 words.* Whether or not such effects interact with proficiency effects remains unclear.

A study by Barbour (1995), involving the monitoring of sentence-embedded words (e.g., 'she watched the *bubbles* floating in the air'), found advanced learners more in tune with frequency effects than intermediate learners, and less so than L1 speakers.

Dejean de la Bâtie and Bradley (1995) gave less consistent results. Students of French had to rate the frequencies of French nouns like 'théâtre' and detect a word-initial target [t] inside sequences like 'grand théâtre' [gʁãteatʁ] ('big theatre'). For first-year students, frequency ratings correlated with target detections (hence, with segmentation of the word starting in [t]). Second-year students, however, showed no such correlation.

A further, tangential indication that familiarity effects do not necessarily increase with L2 proficiency is the lack of clarity as to whether speech segmentation in an unfamiliar language gains from growing lexical exposure (Goldstein, 1983; Vogel & Winitz, 1989; Cowan, 1991) or not (Wakefield, Doughtie & Yom, 1974; Pilon, 1981).

These findings have left unresolved the issue of whether learners relied on lexical frequency cues late (Subhypothesis 3a: shift from phonotactic to lexical cues) or early, and increasingly (Subhypothesis 3b: increased mix of lexical and phonotactic cues).

2.2.1.4. Lexical neighborhood effects on L2 spoken word recognition

*Background: L1 vs. L2 neighborhood.* 'L1 neighborhood' refers here to similarities between so-called 'cognates': L2 words (e.g., English 'lard') that share meaning and phonological or orthographic similarity with L1 words (e.g., French 'lard' [laʁ]). 'L2 neighborhood' refers to similarities between two L2 words (e.g., English 'cap' - 'cab').

*L1 neighborhood effects on the recognition of isolated L2 words.* In the present study of Francophone ESL learners, such effects would translate as effects of Cognacy on lexical decision. One possible outcome is a 'cognate advantage', whereby cognates like 'bus' (in French, 'bus', [bys]) are processed faster and/or more accurately than non-cognates like 'mad' (in French, 'fou', [fu]). The reverse outcome is a 'non-cognate advantage'.

Empirical studies of cognacy effects on L2 spoken word recognition have been sparse and conflicting. Caramazza and Brones (1979) thus found a cognate advantage, whereas Experiment 3 by Sebastián-Gallés, Echeverría and Bosch (2005) did not.

Theories of bilingual lexical processing further hint that cognate and non-cognate advantages are equally plausible to low- and high-proficiency learners. In theoretical terms, a cognate advantage means that the processing of a cognate has not been inhibited by the association of that cognate with an L1 word. This scenario supports the 'selective language view', whereby bilinguals process L2 lexical representations independently from L1 lexical representations (Frenck-Mestre & Prince, 1997; Kroll & Stewart, 1994; Talamas, Kroll, & DuFour, 1999). In contrast, a noncognate advantage means that the processing of a cognate has been inhibited by an L1 word, in accordance with the 'non-selective language view', whereby bilinguals co-process L1 and L2 representations (Cristoffanini, Kirsner, & Milech, 1987; de Bruijn, Dijkstra, Chwilla, & Schriefers, 2001; van Heuven, Dijkstra, & Grainger, 1998). In this light, and regardless of the size of the learners' L2 lexicon (smaller for Intermediate learners, wider for Advanced learners), what should determine whether learners display a cognate or non-cognate advantage is whether or not they can process L2 words without co-activating their L1 lexicon.

Given this ambivalent empirical and theoretical background, no specific prediction could be made as to whether Intermediate or Advanced learners would be more sensitive to a cognate or non-cognate advantage in the L2 lexical decision task.

*L1 neighborhood effects on the recognition/segmentation of string-embedded L2 words.* With the exception of Ju and Luce (2004), eye-tracking studies concerned with such L1 neighborhood effects have found these effects to be inhibitory (Spivey & Marian, 1999; Marian & Spivey, 1999, 2003; Marian, Spivey & Hirsch, 2003; Weber & Cutler, 2004). Marian et al. (2003), in particular, found that, when asked in English to "pick up the marker", Russian-English bilinguals would often glance at the distracting picture of a stamp, the Russian name of which is 'marka' (i.e., an L1 neighbor of English 'marker'). Such findings inspired the hypothesis that Cognacy might affect L2 word-spotting/ segmentation (Hypothesis 2).

*L2 neighborhood effects on the recognition of isolated L2 words.* Phonologically neighboring words like 'sat' and 'set' are often confused by Dutch ESL learners (Swan & Smith, 2001), which can explain why they would misjudge nonwords based on the vowels [æ]-[ɛ] during lexical decisions in English (Broersma, 2002). Along similar lines, Bradlow and Pisoni (1999) found that ESL learners had more trouble transcribing spoken English words if those had many phonological neighbors in L2 English (e.g., 'ban') than if they had few (e.g., 'was'). This conclusion, though challenged by Takayanagi, Dirks and Moshfegh (2002), was supported by Imai, Walley and Flege (2005) (in the condition where L2 English words were not produced with an L1 Spanish accent). Altogether, these suggestions that L2 neighborhood effects might overall be inhibitory further called for the control of neighborhood densities in the lexical decision task of this study.

*L2 neighborhood effects on the recognition/segmentation of string-embedded L2 words.* Preliminary evidence for such effects came from an eye-tracking study by Weber and Cutler (2004). Dutch ESL learners were more distracted by pictures with names evoking highly confusable vowels like [æ]-[ɛ] (e.g., 'pencil', given the target 'panda') than by pictures with names evoking less confusable vowels like [i]-[ɔ] (e.g., 'beetle', given the target 'bottle'). This finding offered a further incentive to control the neighborhood densities of the stimulus English words in the L2 word-spotting task of the present study.

2.2.1.5. Amount of lexical input needed to recognize L2 words

Studies using word monitoring (Goldstein, 1983), phoneme monitoring (Matter, 1986), comprehension of noise-filled sentences (Bond, Moore & Gable, 1996), or gating (Nooteboom & Truin, 1980; Koster, 1987; Chesneau, 1992; Pearman, 2004) show that, in comparison with L1 listeners, L2 listeners need to hear more of a word in order to identify it, and that more proficient L2 listeners more easily detect words before they end. Hence some further motivation for Hypothesis 3 of this study (whereby proficiency may affect L2 segmentation) and for the use of string-final targets in word-spotting.

2.2.1.6. Summary

This discussion has shown the importance of words in L2 speech processing. The above literature has also provided the basis for exploratory Hypotheses 2 and 3 of this study (i.e., effects of Cognacy and Proficiency on L2 word-spotting/segmentation), as well as for two methodological decisions: 1) make word-spotting targets string-final; 2) balance targets for neighborhood densities.

**2.2.2. Sublexical factors**

2.2.2.1. The relevance of sublexical cues to L2 speech segmentation

The multiple-cue approach, advocated earlier with regard to L1 listeners, seems all the more relevant to L2 listeners. L2 listeners have indeed three good reasons to combine lexical segmentation cues with a synergy of diverse sublexical cues. First, they may continuously have to compensate for lexical deficiencies in their target language. Second, because they are constantly pressed to decipher the utterances of their L1 interlocutors, they seem directly affected by the need for efficient speech processing. Third, there are two preliminary pieces of evidence that they can integrate lexical and sublexical information: 1) retrospective evidence that they can detect prominent prosodic material in a sentence only if it appears on a familiar word (Kim, 1995); 2) experimental evidence that they may exploit sublexical-prosodic cues even when the input contains sufficient lexico-semantic information (Sanders, Neville & Woldorff, 2002).

2.2.2.2. Sublexical-phonetic cues

*Phonetic effects on the recognition of isolated L2 words*. Research on the phonetics of L2 speech perception has focused on the perceptual categorization of L2 speech sounds (see, e.g., Major, 2001 for review), so that studies on phonetic cues to L2 spoken word recognition have been comparatively sparse. Some of these studies have found that performance in L2 word recognition correlated positively with the L2 learner's relative knowledge of the relevant word-internal segments (Nábelěck & Donahue, 1984; Koster, 1987; Mayo, Florentine & Buus, 1997). Others have found that, as L2 proficiency grows, learners became more native-like in perceiving the durational variations of word-internal segments (Enomoto, 1992), and that target words produced with no L1 accent were recognized with more success (Imai, Walley & Flege, 2005). These conclusions confirm the potential role of L2 proficiency in the recognition of isolated L2 words.

*Phonetic effects on the recognition/segmentation of string-embedded L2 words.* Bannert, Nicolas and Stridfeldt (1996) found that neither L2 or L1 speakers of French could consistently disambiguate resyllabified French phrases by using phonetic cues. Altenberg (2005) found that Spanish ESL learners, even at intermediate or advanced levels, were significantly worse than natives at using allophonic cues for distinguishing phrases like 'keep sparking' and 'keeps parking' (e.g., the aspirated [p] of 'parking'). Sinor (2002) reported a transcription task where beginning and advanced learners of French had trouble differentiating and segmenting quasi-homophonous phrases like 'petit ami' ('boyfriend'), 'petite amie' ('boyfriend') and 'petit tamis' ('little sieve'). This finding was paradoxical in that the length of the word boundary [t] had been enhanced to yield a cue that would consistently disambiguate such phrases[8]. So far, in light of these studies, it seems that L2 listeners may not gain much from phonetic cues to L2 segmentation.

#### 2.2.2.3. Sublexical-prosodic cues

*Prosodic effects on the recognition of isolated L2 words.* Research into such effects have focused on how L1-based prosodic strategies might interfere with L2 word recognition.

At the origin of this research is a monolingual study by Mehler, Dommergues, Frauenfelder and Ségui (1981). They found a "syllable effect" whereby L1 Francophones would detect a target (e.g., 'pa' [pa]) faster if it matched the first syllable of a stimulus word (as in 'palace' [pa.las], French for 'palace') than otherwise (as in 'palmier'

[8]However, as we have already mentioned (Section 2.1.3.2, footnote 4), this cue did facilitate speech segmentation for all listeners when it was enhanced to simulate a [t] of geminate quality such as the [t] of 'petite Tammy' ('little Tammy'), thus contrasting with the singleton [t] of the above examples.

[pal.mie], French for ‘palm tree’). This syllable effect was later replicated with non-native listeners. Cutler, Mehler, Norris and Ségui (1986), having added English materials to Mehler et al.’s materials, found that Francophones showed a significant syllable effect in English as well as in French. Francophones thus seem to ‘syllabify’ in any input language, even Japanese, which calls for moraic rather than syllabic segmentation (Otake, Hatano, Cutler & Mehler, 1993). Japanese listeners, conversely, use moraic segmentation not just in their language (Otake et al., 1993) but even in English (Cutler & Otake, 1994), French and Spanish (Otake, Hatano & Yoneyama, 1996).

Other studies, however, undermined the risk that L1-based prosodic strategies might interfere with L2 word recognition. Cooper, Cutler and Wales (2002) found that proficient Dutch ESL learners, like native Anglophones, detected visually presented target words faster upon hearing stress-matching word fragment primes (e.g., mus-, from MUSic, or admi- from ADMIral). Cutler, Mehler, Norris and Ségui (1992) also found that L2 listeners could override the interference of L1 prosody. Using the procedure and items from Cutler et al. (1986), they identified two types of French-English bilinguals: 1) the English-dominant ones, who processed both English and French materials ‘the English way’ (i.e., with no sign of syllabic segmentation); 2) the French-dominant ones, who processed French materials ‘the French way’ (i.e., via syllabic segmentation) but were also flexible enough to process English materials without syllabic segmentation.

*Prosodic effects on the recognition/segmentation of string-embedded L2 words.* Some studies suggest that L1 interference makes L2 prosodic segmentation difficult to learn (Broselow, 1988; Cutler, Mehler, Norris & Segui, 1989, 1992; Golato, 2002). In the third experiment by Cutler et al (1992), English-dominant bilinguals had spotted ‘mint’ with more ease if its carrier had first-syllable stress pattern (e.g., ‘mintef’ [’mɪntəf]) than otherwise (e.g., ‘mintayf’ [mɪn’tɛjf]). But French-dominant bilinguals (i.e., highly advanced L2 listeners) did not display this English segmentation strategy.

Other studies suggest that L2 prosodic strategies are in fact learnable (Bradley, Sánchez-Casas & Garciá-Albea, 1993; Goetry & Kolinsky, 2000; Sanders, Neville & Woldorff, 2002). Sanders et al. thus found that ESL learners (beginning and advanced, L1 Spanish and L1 Japanese) detected target phonemes with more ease in words with first-syllable stress (e.g., [b] in 'in order to recycle bottles you have to separate them').

2.2.2.4. Sublexical-phonotactic cues

So far, there seems to be no experimental evidence for the influence of phonotactic factors on the recognition of isolated L2 words. Below, however, is some preliminary evidence for phonotactic effects on the recognition/segmentation of string-embedded L2 words. This issue is central to the present study since Onset Type acts as an independent variable in the forthcoming word-spotting experiment (Experiment 4).

Pioneering studies by Andrea Weber and colleagues (Weber, 2000; Weber & Cutler, 2004, 2006) offer some indication that knowledge of L2 phonotactics can influence the speech segmentation performance of L2 learners, at least those with an advanced level of L2 proficiency. In focus here is the latest of these studies.

German speakers of high L2 English proficiency and native Anglophone controls were asked to spot target English words inside nonsense carrier strings. The experiment explored the four contrastive conditions offered by English and German for the alignment or misalignment of word and syllable boundaries (Table 2)[9].

Table 2. Summary of phonotactic stimulus conditions in Weber and Cutler (2006)

| Cross-linguistic condition (E: English; G: German) | Carrier string for the target word 'lance' | String-medial CC cluster of interest | Cluster's status in L2 English (E) & L1 German (G) |
|---|---|---|---|
| 1. Aligned in E and G | [dʒimlæns] | [ml] onset | illegal in E and G |
| 2. Aligned in E, Misaligned in G | [ðiʃlæns] | [ʃl] onset | illegal in E, legal in G |
| 3. Misaligned in E, Aligned in G | [blɔislæns] | [sl] onset | legal in E, illegal in G |
| 4. Misaligned in both E and G | [θiplæns] | [pl] onset | legal in both E and G |

[9]For explanatory purposes, these conditions were renamed using the Aligned and Misaligned terms used so far in this dissertation. Conditions 1, 2, 3 and 4, as summarized in this table, were respectively called 'Common boundary', 'English boundary', 'German boundary' and 'No boundary' in Weber and Cutler's original report.

The first key finding, reached using Condition 4 (Misalignment in English and German) as a baseline, was that word-spotting was faster and more accurate under both Condition 1 (Alignment in both German and English) and Condition 2 (Alignment in English only). Advanced L2 learners can thus exploit L2 phonotactic cues specific to their L2. This finding inspired Hypothesis 1 of the present study, whereby L2 phonotactic cues, along with L2 lexical cues, may affect L2 speech segmentation.

The second key finding was that word-spotting was faster and more accurate in Condition 3 (Alignment in German only). At first glance, this suggests that L1 phonotactics can help L2 segmentation. On second consideration, it means that German listeners were using L1 cues of no relevance to segmentation in L2 English. Thus, in German, the [sl] cluster is an illegal onset, conducive to facilitating Alignment effects; but in L2 English, it has the opposite status, and so should have caused inhibiting rather than the observed facilitating effects. Take, also, the clusters [sw] and [tw], used in Condition 3 to elicit the segmentation of w-starting words. Before submitting these clusters to the constraint 'illegal onset entails Alignment' (as they seem to have done), listeners would have had to assimilate the L2 consonant [w] to the L1 consonant [v]. This can be seen as yet another case of L1 interference, hence further motivation for Hypothesis 2 of the present study, whereby L1 phonotactic/Onset Type factors might affect L2 segmentation.

The conclusions of this study are that: 1) L2 listeners can learn to use L2 phonotactic segmentation cues efficiently; 2) they remain vulnerable to the effects of L1 phonotactics. Put together, these two conclusions mean that the importance of L1 transfer on the application of L2 phonotactic cues remains unresolved (as was the case with L2 prosodic cues – cf. Section 2.2.2.3). One of the objectives of the present study is to readdress this issue. Another is to assess how the use of L2 phonotactic cues may evolve over the course of L2 acquisition. The need to address this developmental issue was expressed by Weber and Cutler (2006: 604) themselves: "These listeners, it is true, were very proficient indeed in their L2; we cannot say on the basis of the present results exactly how much L2 experience is necessary for the efficient exploitation of L2 boundary constraints which do not apply in the L1".

2.2.2.5. Summary

These studies of sublexical cues to L2 word processing have motivated and/or specified the three exploratory, two-tailed hypotheses of the present study: Hypothesis 1 (i.e., the possibility that L2 phonotactic/Onset Type cues might contribute to L2 segmentation along with L2 lexical/Frequency cues); Hypothesis 2 (the possibility that L1 phonotactic/Onset Type factors might affect L2 segmentation – and not necessarily negatively); Hypothesis 3 (the possibility, so far largely overlooked, that the use of L2 phonotactic segmentation cues might evolve of the time course of L2 acquisition).

### 2.2.3. Conclusion

This review of the L2 literature has confirmed the possibility that L2 listeners might segment L2 input via a combination of lexical and sublexical cues. At the same time, empirical issues pertaining to the helpfulness of L2 lexical and sublexical cues remain wide open. In this open empirical context, here again is the aim of the study reported next: to assess the contributions of L1 and L2 lexical and phonotactic cues to the performance and the acquisition of L2 speech segmentation.

# CHAPTER THREE

# PILOT TEST OF PHONOTACTIC CUES

## 3.1. RATIONALE

A first step in testing the role of phonotactics in L2 speech segmentation is to choose CC clusters capable of offering phonotactic cues to L2 listeners. Specifically, to test the possibility of L1 phonotactic effects, one needs clusters of comparable or contrastive phonotactic statuses in the L1 and the L2. In the present case, this task was complicated by the researcher's bilingual intuition that English and French phonotactics are highly resemblant. It was decided, therefore, that a distributional analysis of CC clusters in English and French (Section 3.2), followed by off-line tests of cluster segmentation/acceptability by native Anglophones (Sections 3.3 and 3.4), would help select these clusters more precisely and reliably than would bilingual intuition alone[10].

## 3.2. CORPUS-BASED COMPARISON OF ENGLISH AND FRENCH PHONOTACTICS

*Step 1. Corpus preparation.* A distributional analysis of English and French CC clusters began with the selection of a subset of 1092 monosyllabic English words from the MRC Psycholinguistic Database (Coltheart, 1981) as well as of a subset of 646 monosyllabic French words from the Lexop database (Peereman & Content, 1999). Words were selected to appear in these two corpora if they had a CCVC or CVCC structure and if they were made of phonetic segments common to English and French[11].

[10]Frequencies of cluster distribution could have been borrowed from external sources such as: for French, Wioland (1985); for English, Buchwald (2005), who compiled frequency data using the Celex database (Baayen, Piepenbrock and Gulikers, 1995). However, a hands-on comparison of English and French phonotactics seemed to enhance the validity of this pilot study, in that it made necessary to compute frequency counts using exactly the same counting method for English and French (see Section 3.2, Step 2).

[11]Lexop uses monosyllabic words from the Brulex database (Content, Mousty & Radeau, 1990), which is itself made of entries from the Micro Robert dictionary (Robert, 1986). MRC uses words from printed and spoken sources such as the Edinburgh Associative Thesaurus (Kiss et al, 1973) and the Cornell University tape of 20,000 commonly used words. This mix of spoken and written sources was not expected to alter the cluster counts dramatically. As Buchwald (2005: 74) points out, it is "not clear what systematic differences should be expected when using both spoken and written corpora compared to just using a spoken corpus".

*Step 2. Computation of cluster frequencies.* For both English and French, cluster frequencies were computed as the average percentage frequencies of the CC onsets and CC offsets found among the words of the English and French corpora (see Appendix A).

*Step 3. Phonotactic assessment of English and French CC clusters.* While these frequency data promised an easy way to set up phonotactic cues of relative strengths in the L2 segmentation experiments, identifying relative cue strengths a priori required some arbitrary choices for maximizing the information offered by these data.

The first choice, following customary practice, was to treat rare clusters as accidental gaps in the phonotactics of English and to regroup them alongside illegal clusters. The upper frequency value for illegal/accidental clusters was set at 0.2% because, in 3 out of 4 of the left-skewed distributions of English and French clusters (i.e., the distributions of English onsets, English offsets and French offsets), 0.2% was the first frequency value to be found between 0% and the lowest modal peaks.

The second choice, with respect to legal/non-accidental clusters, was to set at 3% the cut-off point between low and high frequencies. 3% was indeed simultaneously equal to the mean and median frequency of CC clusters in the English corpus.

The third choice, with regard to cluster distributions, was to equate the terms 'low vs. high frequency' with the terms 'marked vs. unmarked', so as to facilitate later result discussions that might consider both the variables Onset Type and Frequency.

These choices and the resulting quantitative criteria for the phonotactic assessment of CC clusters are summarized in Table 3.

Table 3. Quantitative criteria for the phonotactic assessment of CC clusters

| Phonotactic status | Frequency range | Example | |
|---|---|---|---|
| | | in English | in French |
| illegal/accidental | 0% - 0.2% | [sf] onset, [ln] offset | [mp] offset |
| marked | 0.3% - 3% | [gl] onset, [kt] offset | [sp] onset, [ks] offset |
| unmarked | > 3% | [tr] onset, [mp] offset | [kl] onset, [sk] offset |

*Step 4. Comparison of cluster frequencies between English and French.* These criteria allowed the identification of 5 main groups of consonant clusters that were homogeneous in terms of both phonetic content and cross-lingual phonotactic status (see Table 4).

Table 4. Comparative phonotactic statuses of CC clusters in English and French

| CC clusters (types & tokens) | | Frequencies per language English (E) | French (F) | Cross-lingual phonotactic status |
|---|---|---|---|---|
| <u>Onsets</u> | | | | |
| *stop+[l] | [bl, fl, gl, kl, pl] | 4.4% | 5.0% | unmarked in E, unmarked in F |
| *stop+[r/ʁ] | [br/bʁ, fr/fʁ, gr/gʁ, kr/kʁ, pr/pʁ, tr/tʁ] | 6.0% | 9.6% | unmarked in E, unmarked in F |
| *[s]+stop (except [st], which is unmarked in French) | [sk, sm, sn, sp] | 4.6% | 1.2% | unmarked in E, marked in F |
| <u>Offsets</u> | | | | |
| *[l]+obstruent (except [ld/lt]) | [lf, lk, lm, lp] | 2.7% | 1.2% | marked in E, marked in F |
| *[m/n]+obstruent (except [mf]) | [mp, nd, nk, ns, nt] | 6.0% | 0.0% | unmarked in E, illegal in F |

*Step 5. Implications for the main study.* Onset clusters that are unmarked in English but marked in French (i.e., [sk, sm, sn, sp]) could result in phonotactic cues of contrastive weights from the participants' L1 and L2. Meanwhile, onsets that are unmarked in English and French (i.e., [bl, fl, gl, kl, pl] and [br/bʁ, fr/fʁ, gr/gʁ, kr/kʁ, pr/pʁ, tr/tʁ]) could act as controls for testing L1 transfer effects. But before testing whether L2 segmentation could be affected by subtle cross-lingual variations in the markedness of legal CC clusters, a more basic claim needed to be tested: whether the actual legality or illegality of a given cluster could affect the use of phonotactic cues in segmentation per se. This test was run with native speakers of English in Experiments 1 and 2, as reported next.

## 3.3. EXPERIMENT 1: OFF-LINE SEGMENTATION OF PRINTED CVCCCVCs IN L1 ENGLISH

### 3.3.1. Rationale

Three aims drove Experiment 1 (as well as Experiment 2, its auditory replica):
1) to assess the extent to which Anglophones segmented CVCCCVC strings by assessing the phonotactic statuses of overlapping, string-medial CC clusters;
2) to determine whether CC onsets or CC offsets would better elicit phonotactic effects in the subsequent, main segmentation experiments with French ESL learners;
3) to finalize the format of stimuli for these main experiments.

### 3.3.2. Participants

Sixty undergraduate students from the University of Alberta were recruited during an introductory linguistics class. They were presented with the topic of the present research before being given the printed sheets needed for this offline task.

Forty of these students, having actually filled in the sheets and indicated that they were native Anglophones, qualified as voluntary participants in the experiment.

### 3.3.3. Materials

Sixty-eight CVCCCVC nonsense strings (see Appendix B) were drawn from the ARC Nonword Database (Rastle, Harrington & Coltheart, 2002) so that their CCC portions were pronounceable, unfound in English, and made of overlapping CC clusters characteristic of one of the following phonotactic categories:
1) LEG-LEG (i.e., 'legal offset - legal onset', as in 13 items like 've<u>nkl</u>of');
2) LEG-ILLEG (i.e., 'legal offset - illegal onset', as in 23 items like 'bi<u>mpn</u>on');
3) ILLEG-LEG (i.e., 'illegal offset - legal onset', as in 23 items like 'sa<u>fbl</u>eg');
4) ILLEG-ILLEG (i.e., 'illegal offset - illegal onset', as in 9 items like 'noo<u>tkp</u>um').

The CC clusters were selected in light of the analysis of CC cluster distribution. For the LEG-LEG items, the legal offsets and onsets were chosen to be statistically unmarked, so that they would have comparable chances of being preserved in the participants' transcriptions. For the ILLEG-ILLEG items, 'accidental clusters' (i.e., clusters with a low frequency of 0.1% or 0.2%) were included along with full-fledged 'illegal clusters' (i.e., clusters with 0% frequency). The idea was to test whether minor variations in sonority and cluster frequency might interact during the segmentation of ILLEG-ILLEGs (Section 3.2.4.4).

A subset of the critical CC clusters took the shape of duplicate, geminate-like consonants (as in 'nebbluk'), to reflect the possibility that gemination might be a source of segmentation cues (see Section 2.1.2.2). In Experiment 2, auditory responses to stimuli like 'nebbluk' would help decide whether to insert geminate cues in the main segmentation tasks. These stimuli were hence called 'strings with geminates'.

### 3.3.4. Procedure

The stimulus strings were shown on a printed sheet. For each string, participants were asked to indicate their favored segmentation by drawing a slash in the desired position, as in 'femt/kas' or 'fem/tkas'. No other transcription format was allowed (nor found among the responses). Two random orders of stimulus presentation were assigned.

### 3.3.5. Results and discussion

#### 3.3.5.1. Data coding and data distribution

Transcriptions were coded CC-C (e.g., 'lunk/min') or C-CC (e.g., 'fik/flaf'). Table 5 shows the distribution of these transcription types.

Table 5. Distribution of transcription types per category of printed CVCCCVC strings

| | ILLEG-LEG strings | LEG-ILLEG strings | LEG-LEG strings | ILLEG-ILLEG strings |
|---|---|---|---|---|
| CC-C responses | 7% | 98% | 17% | 67% |
| C-CC responses | 93% | 2% | 83% | 33% |

3.3.5.2. Responses to fully and partially legal strings

*Main patterns.* Participants appeared to know their English phonotactics, in light of the first two columns of Table 5. For ILLEG-LEG strings, they accurately chose C-CC (i.e., onset-preserving) responses over CC-C (i.e., offset-preserving) responses. For LEG-ILLEG strings, they accurately chose CC-C over C-CC responses.

Participants also appeared to prefer legal onsets over legal offsets. Thus, while onsets and offsets of LEG-LEG items are both legal, hence both worth preserving, the third column of Table 5 shows that participants largely preferred the C-CC rather than CC-C layout. This pattern reflects the typologically attested bias for onset maximization (Blevins, 1995; Prince & Smolensky, 1993).

*Test of graphemic influences.* The statistical data for this test (as reported item-by-item in Appendix C), were: 1) a cluster preservation rate (i.e., for each legal onset/offset, the percentage of onset- or offset-preserving transcriptions among all the relevant stimulus transcriptions); 2) n-gram frequencies, taken from Baayen, Piepenbrock and Gulikers's (1995) Celex lexical database (i.e., the frequencies at which clusters of consonant graphemes overlapping the CC onsets or offsets of interest appear in any position in printed English words of any length, per million of English words).

The rationale for this test was that the segmentation of a legal CC onset like [kl], in 'venklof', could have been influenced by the frequency (hence relative admissibility) of at least one of three grapheme clusters: the trigram nkl-, the offset bigram nk-, the onset bigram kl-. The same reasoning applied to legal CC offsets like [nt], in 'muntpin', possibly influenceable by the trigram ntp-, the offset bigram nt-, the onset bigram tp-.

Actual testing of the relationship between n-gram frequencies and preservation rates used Pearson product-moment correlation coefficients. The results of this test (see Table 6) indicate that graphemic frequencies had in fact not played a significant role in the participants' segmentations of fully and partially legal strings.

Table 6. Correlations of n-gram frequencies with the preservation of printed clusters

| | Onset preservation rate | Offset preservation rate |
|---|---|---|
| offset bigram frequency | r(9) = -0.44, p = 0.23 | r(7) = 0.29, p = 0.53 |
| onset bigram frequency | r(9) = -0.11, p = 0.77 | r(7) = -0.16, p = 0.73 |
| trigram frequency | r(9) = 0.25, p = 0.51 | r(7) = -0.35, p = 0.44 |

*Strings with geminates.* The relevant ILLEG-LEGs and LEG-ILLEGs respectively got 87% C-CC responses (e.g., 'neb/bluk') and 98% CC-C responses (e.g., 'sant/tib'). Responses like ne/bbluk or san/ttib were thus largely avoided, which is not surprising since English words may not start with double consonant graphemes. Experiment 2 would still have to retest these stimuli for gemination effects on auditory segmentation.

3.3.5.3. Responses to illegal strings

*Main patterns.* While ILLEG-ILLEGs received a 67% preference for CC-C responses (Table 5), 4 of these items got equal proportions of CC-C and C-CC responses (Table 7).

Table 7. Distribution of transcriptions for the printed ILLEG-ILLEG strings

| | ILLEG-ILLEG string | Transcription Type (%) | |
|---|---|---|---|
| | | CC-C | C-CC |
| 1. | femtkas | 100 | 0 |
| 2. | nimkfon | 98 | 20 |
| 3. | minfpom | 93 | 7 |
| 4. | papfken | 68 | 32 |
| 5. | fotfsut | 58 | 42 |
| 6. | mekpfap | 50 | 50 |
| 7. | nootkpum | 50 | 50 |
| 8. | pekfmip | 50 | 50 |
| 9. | tetpsop | 50 | 50 |

A possible explanation for this variability of responses is that listeners had processed each ILLEG-ILLEG item by assessing the sonority and frequency of the relevant illegal/accidental CC clusters. This phonotactic hypothesis is explored next.

*Phonotactic analysis.* This analysis invokes the scale of increased sonority across phonetic segments, or 'Sonority Scale': Plosives–Fricatives/Affricates–Nasals–Liquids–Glides–Vowels (Selkirk, 1984). Also relevant in this analysis are two principles usually thought to dictate the shape of syllable onsets and offsets in English.

The first of these principles, Sonority Sequencing (SS), holds that sonority must increase gradually from the onset to the peak of a syllable, and decrease gradually from peak to offset (Selkirk, 1984). SS is fulfilled by the onset cluster [pr] (as well as by all English onset CC clusters other than '[s]+obstruent' clusters). SS, however, is violated by the [pf] offset (as the succession of a plosive and a fricative constitutes an increase of sonority rather than the expected decrease of sonority at the end of a syllable).

The next principle, Minimal Sonority Distance (MSD), holds that the distance between the members of a CC cluster must be of two degrees or more on the Sonority Scale (Selkirk, 1984; Clements, 1990). MSD is fulfilled by the onset cluster [fl] (as fricatives and liquids are two degrees apart on the above scale). MSD, however, is violated by the [pf] onset (as plosives and fricatives are one degree of sonority apart).

In this light, ILLEG-ILLEGs would have given rise to preferred segmentations vs. preferrable segmentations. The preferred segmentations appear in Table 7 above (e.g., 93% CC-C responses to 'minfpom'). The preferrable segmentations, meanwhile, would favor at least one of two scenarios: 1) a cluster that made fewer sonority violations than an adjacent cluster; 2) an accidental cluster with a slightly higher usage frequency than an adjacent, accidental cluster. Thus, the segmentation 'fe<u>mt</u>/<u>k</u>as' seems preferrable to 'fe<u>m</u>/<u>tk</u>as' because the [mt] offset is itself preferrable: 1) it obeys SS and MSD, whereas the [tk] onset violates MSD; 2) it has quasi-null frequency, whereas the [tk] onset has a zero-frequency [tk]. In this light, Table 8 sums up the comparisons of preferrable and preferred segmentations.

Table 8. Sonority and frequency in the segmentation of printed ILLEG-ILLEG strings

| Stimulus | Status of Offset Cluster | | Status of Onset Cluster | | Corresponding Segmentation Pattern | |
|---|---|---|---|---|---|---|
| | Sonority | Fcy | Sonority | Fcy | Preferrable | Preferred |
| femtkas | Tolerable | Low | Bad* | Zero | CC-C | CC-C (100%) |
| nimkfon | Tolerable | Low | Bad* | Zero | CC-C | CC-C (98%) |
| minfpom | Bad* | Low | Very bad** | Zero | CC-C | CC-C (93%) |
| papfken | Very bad** | Low | Very bad** | Zero | CC-C | CC-C (68%) |
| fotfsut | Very bad** | Zero | Bad* | Zero | C-CC | CC-C (58%) |
| mekpfap | Bad* | Zero | Bad* | Low | CC-C | CC-C (50%), C-CC (50%) |
| nootkpum | Bad* | Zero | Bad* | Zero | CC-C/C-CC | CC-C (50%), C-CC (50%) |
| pekfmip | Very bad** | Zero | Bad* | Zero | C-CC | CC-C (50%), C-CC (50%) |
| tetpsop | Bad* | Zero | Bad* | Zero | CC-C/C-CC | CC-C (50%), C-CC (50%) |

*Violation of the MSD Principle only **Violation of both the SS and MSD Principles

Note. SS violations refer here to clusters whose sonority varies in an illicit direction along the Sonority Scale (e.g., [fp] onset in 'minfpom'), but not to clusters whose elements are of equal degree on the scale (e.g., [tk] onset in 'femtkas').

Table 8 reveals an overall match between preferred and preferrable segmentations, as if ILLEG-ILLEGs had been segmented along the following algorithm:

1) "if adjacent illegal/accidental clusters make unequal numbers of sonority violations, preserve the one that makes fewer sonority violations (as in 'femt/kas')";

2) "if adjacent illegal/accidental clusters make equals number of sonority violations, preserve the one that occurs slightly more frequently in English (as in 'papf/ken')";

3) "if adjacent illegal/accidental clusters make the same number of sonority violations and share zero-frequency, preserve either cluster (e.g., 'nootk/pum', 'noot/kpum')".

Given this hypothetical algorithm and the supporting behavioral data reported above, native Anglophones appear to be phonotactically driven. They seem indeed likely to use sonority constraints and, to a lesser extent, cluster occurrence frequencies, when processing consonant clusters that at first glance seem impossible in English.

### 3.3.6. Summary

Experiment 1 has confirmed the relevance of phonotactics to L1 segmentation. More specifically, it has brought evidence that Anglophones segmented printed CVCCCVC strings by means of three strategies: 1) preserve legal onsets and offsets; 2) prefer legal onsets over legal offsets; 3) tolerate accidental onsets/offsets if they fulfill sonority constraints and/or if they have a low- rather than zero-occurrence frequency.

These strategies were highlighted using printed items. This methodology allowed the prevention of mishearings of the critical clusters, and thus had the merit of yielding transparent insights into cluster acceptability in English. Experiment 2, as reported next, uses the same stimuli to test whether the above strategies extend to the auditory mode.

## 3.4. Experiment 2. Off-line segmentation of spoken CVCCCVCs in L1 English

### 3.4.1. Rationale

This experiment, being an auditory replica of Experiment 1, followed the rationale presented in Section 3.3.1.

### 3.4.2. Participants

Sixty-four native Anglophones took part, all recruited among the undergraduate linguistics students from the University of Alberta.

### 3.4.3. Materials

The materials and stimulus categories were the same as for Experiment 1.

### 3.4.4. Procedure

Materials were read by the researcher with flat intonation and regular pace. String-medial consonants were lengthened for strings with geminates (e.g., 'dinkkep').

Participants wrote each stimulus on a blank form using Roman alphabet (as in a dictation) and indicated their favored segmentation with a space in the desired position, as in 'femt kas' or 'fem tkas'. These two transcription schemes were practiced using 8 separate items, split equally between stimulus categories. But these transcription schemes were not imposed, so that possibly erroneous transcriptions (e.g., 'fe mtkas' or 'fet kmas') might reveal segmentation difficulties in specific phonotactic environments.

Two main sessions were run. In the first, 49 participants heard the randomized LEG-ILLEG, ILLEG-LEG and LEG-LEG items. In the second, 15 more participants heard the randomized ILLEG-ILLEG items. This second, separate session helped prevent fatigue effects. It also ensures that responses to fully illegal items were not affected by the cluster-building strategies underlying responses to fully or partially legal items.

### 3.4.5. Results and discussion

3.4.5.1. Data coding and data distribution

Transcriptions were coded CC-C (e.g., 'lunk min'), C-CC (e.g., 'fik flaf'), or 'Other' if they showed no space (e.g., fikflaf'), a disformed or unintended cluster (e.g., 'fib laf', 'fik plaf'), a CCC offset (e.g., 'menkt es'), or metathesis (e.g., 'fif klaf').

Distribution of these three transcription types (see Table 9) is interpreted next.

Table 9. Distribution of transcription types per category of spoken CVCCCVC strings

| | ILLEG-LEG strings | LEG-LEG strings | LEG-ILLEG strings | ILLEG-ILLEG strings |
|---|---|---|---|---|
| CC-C responses | 1% | 22% | 29% | 36% |
| C-CC responses | 73% | 61% | 0% | 15% |
| Other responses | 26% | 17% | 71% | 49% |

3.4.5.2. Responses to fully and partially legal strings

*Main patterns.* An accurate bias for legal onsets reemerges, with more onset-preserving responses (C-CC) than onset-violating responses (CC-C or Other) to ILLEG-LEGs.

Less clear is the participants' readiness to preserve legal offsets when needed. While expectingly receiving no C-CC responses, LEG-ILLEGs only got 29% offset-preserving (CC-C) responses and 71% Other responses. These last are discussed below.

*Subanalysis of Other responses.* Proportions of Other responses were expected to be:

a) lowest for LEG-LEGs (as two legal clusters should boost C-CC or CC-C responses);

b) low for ILLEG-LEGs (as an illegal offset present a risk for confusion);

c) high for LEG-ILLEGs (as the principle of onset maximization should make an illegal onset even more confusing than an illegal offset);

d) highest for ILLEG-ILLEGs (as fully illegal strings present the worst-case scenario).

Hence the expected ranking LEG-LEG < ILLEG-LEG < LEG-ILLEG < ILLEG-ILLEG. The obtained ranking was LEG-LEG < ILLEG-LEG < ILLEG-ILLEG < LEG-ILLEG.

Two reasons can explain why LEG-ILLEG finally got most Other responses.

First, LEG-ILLEGs (which required the preservation of legal offsets) could have in fact been more confusing than ILLEG-ILLEGs (which did not require an unnatural preference for illegal over legal clusters).

Second, a closer examination of Other responses to the LEG-ILLEGs highlights two revealing patterns. One portion of Other responses (e.g., 'niskem', for the LEG-ILLEG item [nistkem]) shows that listeners occasionally deleted string-medial segments as if to avoid implementing illegal onset clusters (here, [tk]). Another portion of Other responses (e.g., 'namplis', for the LEG-ILLEG item [nantlis]) shows that listeners occasionally perceived legal onsets that were in fact absent from the LEG-ILLEG items (in this example, [pl]). These two patterns confirm the main trend of Experiment 2, namely a bias for legal onsets.

*Strings with geminates.* LEG-ILLEGs with geminates received 82% Other responses, which by far outweighed the collapsed 1% C-CC and 17% CC-C responses. ILLEG-LEGs with geminates received 51% Other responses, which also outweighed the collapsed 47% C-CC and 2% CC-C responses.

This general bias for Other responses was illustrated by a tendency to transcribe a geminate consonant as two singletons split over a syllable boundary (e.g., 'ne<u>b</u> <u>bl</u>uk'). It seemed that this transcription strategy, while confirming the potential importance of gemination in speech segmentation, could complicate analyses of cluster preservation patterns in subsequent experiments using CVCCVC(C) strings. String-medial geminates seemed therefore best left aside from those main experiments.

### <u>3.4.5.3. Responses to illegal strings</u>

*Main patterns.* Two trends emerge here: items mainly transcribed as CC-C or C-CC vs. items mainly transcribed as Other (see top vs. bottom parts of Table 10).

Table 10. Distribution of transcriptions for the spoken ILLEG-ILLEG strings

| | Stimulus | Transcription Type (%) | | |
|---|---|---|---|---|
| | | CC-C | C-CC | 'Other' |
| 1. | minfpom | 87 | 7 | 6 |
| 2. | papfken | 73 | 7 | 20 |
| 3. | femtkas | 67 | 7 | 26 |
| 4. | pekfmip | 27 | 47 | 26 |
| 5. | nootkpum | 30 | 30 | 40 |
| 6. | nokpfis | 0 | 0 | 100 |
| 7. | fotfsut | 7 | 7 | 86 |
| 8. | mekpfap | 13 | 13 | 74 |
| 9. | tetpsop | 13 | 20 | 67 |
| 10. | nimkfon | 40 | 13 | 47 |

*Phonotactic analysis.* This distribution of CC-C and C-CC responses can again be analyzed in terms of cluster sonority and frequency[12]. From this point of view, a match emerges again between preferrable and preferred segmentations for ILLEG-ILLEGs (see Table 11). This situation confirms the algorithm hypothesized earlier (Section 3.3.5.3), and so the general conclusion drawn from that earlier discussion continues to apply here: sonority and occurrence constraints seem to interact during English speech segmentation, even when a seemingly impossible cluster (e.g., [pf]) is at stake.

Table 11. Sonority and frequency in the segmentation of spoken ILLEG-ILLEG strings

| Stimulus | Status of Offset Cluster | | Status of Onset Cluster | | Corresponding Segmentation Pattern | |
|---|---|---|---|---|---|---|
| | Sonority | Fcy | Sonority | Fcy | Preferrable | Preferred |
| minfpom | Bad | Low | Very bad | Zero | CC-C | CC-C (87%) |
| papfken | Very bad | Low | Very bad | Zero | CC-C | CC-C (73%) |
| femtkas | Tolerable | Low | Bad | Zero | CC-C | CC-C (67%) |
| pekfmip | Very bad | Zero | Bad | Zero | C-CC | C-CC (47%) |
| nootkpum | Bad | Zero | No | Zero | CC-C/C-CC | CC-C (30%), C-CC (30%) |

[12]This analysis ignores ILLEG-ILLEGs primarily transcribed as Other, as such items fail to address whether adjacent illegal/accidental clusters competed on the basis of sonority and/or frequency.

### 3.4.6. Summary

Experiment 2 has brought evidence that Anglophones dealing with spoken CVCCCVC strings tended to use the same phonotactic strategies as when dealing with printed CVCCCVC strings: 1) preserve legal onsets and offsets; 2) prefer legal onsets over legal offsets; 3) tolerate accidental onsets/offsets if they fulfill sonority constraints and, alternatively, if they have a low rather than zero occurrence frequency.

## 3.5. CONCLUSIONS OF THE PILOT STUDY

The findings of auditory pilot Experiment 2, while confirming those of visual pilot Experiment 1, have two implications for the main study.

First, evidence of a consistent bias for onsets over offsets indicates that phonotactic effects on L2 segmentation should be best tested using CC onsets. Thus, in the upcoming word-spotting task, CVCCVC(C) carrier strings like 'zuplaugh' and 'zutlaugh' should facilitate the testing of whether L2 segmentation relies on the comparative statuses of the [pl] and [tl] onsets[13]. Here, therefore, is the chosen format of stimuli for the forthcoming segmentation tasks: CVCCVC(C) strings built on CC clusters from the set [bl, br, fl, fr, gl, gr, kl, kr, pl, pr, sk, sm, sn, sp].

Second, this pilot study gave a hint that L2 word recognition/segmentation could be affected by sonority constraints and/or cluster frequency variations (notably, in light of the cluster distribution analysis, the contrast between the [sk, sm, sn, sp] onsets, marked in French but not in English, and the [bl/br, fl/fr, gl/gr, kl/kr, pl/pr] onsets, unmarked in both French and English[14]).

Chapter 4, which follows next, reports the testing of such effects. The chapter begins by discussing a range of subject- and procedure-related issues that pertain to the multi-experimental study as a whole (Section 4.1).

---

[13] CCC clusters (as in 'zumplaugh' and 'zumtlaugh') would suffice to test the segmentation of English offsets as well (i.e., [mp] and [mt]), as in the pilot. But such materials risk increasing the participants' processing workload and their target detection errors. This risk, already inherent to the word-spotting paradigm, should be minimized when dealing with lexically disfavored L2 learners.

[14] Behavioral evidence for this contrast, and the potential implications of this contrast on English word-spotting by Francophones, are explored in Section 4.3.4.4 in light of a study by Beaudoin (1996).

# CHAPTER FOUR

# MAIN EXPERIMENTS WITH ESL LEARNERS

## 4.1. METHODOLOGICAL PRELIMINARIES

### 4.1.1. Pooling

Twenty-six Chinese ESL learners from Edmonton took part in a pilot run of the lexical decision task (Experiment 3).

Fifty-one Francophone ESL learners from Edmonton and Montréal took part in lexical decision, word-spotting, offline word-likeness judgment and offline segmentation (Experiments 3, 4, 5 and 6). The results of 15 of these Francophone participants were discarded on a post hoc basis. For 3 of these participants, two problems arose: 1) abnormally higher scores in lexical decision and word-spotting than for the rest of the pool; 2) questionnaire evidence that they had been using English daily for over 6 years, thus displaying a profile closer to that of a 'balanced bilingual' than to the 'L2 learner' profile of interest. For 12 more participants, scores were abnormally slow at lexical decision (2 participants) or at word spotting (10 participants), despite detailed task instructions, practice sessions, and the standard request to be as quick and accurate as possible (cf. procedural Sections 4.2.3 for lexical decision and 4.3.3 for word-spotting).

The resulting, final pool of participants comprised 36 Francophone learners of English and 26 Chinese learners of English.

### 4.1.2. L2 proficiency assessment

#### 4.1.2.1. Quantitative assessment

Accuracy scores from lexical decision (Experiment 3) and word-spotting (Experiment 4) were the first means by which L2 proficiency was rated. This method had the advantage of ensuring that the Proficiency variable was closely relevant to the tasks being employed. But this method also prevented from using accuracy scores in later tests of main and interaction effects of Proficiency[15]. Tests of Proficiency effects, in Experiments 3 and 4, would therefore only use reaction time as a dependent variable.

Proficiency scores were computed, for each participant, as an average of lexical decision and word-spotting accuracy scores. Averaging was done by weighing these two scores equally. Indeed, the skills tested in lexical decision and word-spotting (i.e., recognizing isolated and string-embedded L2 words) were deemed equally important parts of the auditory L2 lexical proficiency tested in this study. Proficiency scores (Appendix F) ranged from 45% to 81%. The cut-off point between Intermediate and Advanced scores was set at 64%, a value equal to both the median and average scores.

#### 4.1.2.2. Qualitative assessment

A back-up Proficiency check was run by means of a questionnaire (Appendix E). This questionnaire focused on the participants' educational/professional and ESL learning backgrounds. It was administered before the experiments, orally, and in English.

The first conclusion of this questionnaire is that these ESL learners of English should be split into 'Intermediate vs. Advanced' rather than 'Beginner vs. Advanced' groups. The weakest learners had indeed had 6 years of ESL training before their CEGEP training or had undergone placement tests before attending English university courses.

Intermediate and Advanced learners proved to be similar in age (average age of 23[16]), gender (12 women and 6 men in the Intermediate group; 11 women and 7 men in the Advanced group), occupation (15 students and 3 full-time employees in both groups).

Intermediate participants comprised 2 participants with no more than 3 years of ESL training, and 16 who had only used ESL in secondary school. The 3 Intermediate full-time workers (viz., military, chemistry research assistant, clerk in a French organization) spoke no English at work. The 15 Intermediate students comprised 9

[15]The word-spotting accuracy scores used for Proficiency assessment had in fact been obtained via a different, more liberal, scoring procedure from the word-spotting accuracy scores put under test in Experiment 4 (Section 4.3.4.1). But, despite these differences in scoring procedures, it felt safer to ignore word-spotting accuracy scores from Proficiency tests altogether.

[16]The other age similarities were: a standard deviation of 7 years for the Intermediate and of 6 years for the Advanced; an age range of 16-44 for the Intermediate and 17-37 for the Advanced; a proportion of 5 participants over age 25 among the Intermediate and of 4 among the Advanced.

CEGEP[17] students having just returned to ESL classes (no less than 2 years after their CEGEP training in L2 English) and 6 students in beginner ESL courses (at private language schools). Finally, Intermediate participants tended to rate themselves as not very active in their use of the L2 and as average or poor L2 listeners.

Advanced participants comprised: 6 participants who had been using English daily for at least one year; 9 who had frequented English settings since childhood; 1 who had been perfecting his school training with three extra years of Advanced ESL classes[18]. The 3 Advanced, full-time workers worked in Anglophone companies. The 15 Advanced students included 3 PhD or MBA students, 8 undergraduate students taking Anglophone courses other than remedial/ESL courses at the University of Alberta; and 2 undergraduate students taking advanced ESL courses at the University of Alberta[19]. Finally, Advanced participants tended to rate themselves as active in their use of the L2 and as good L2 listeners.

In sum, learners that had been labelled Intermediate and Advanced on the basis of accuracy scores displayed qualitative profiles that also seemed worthy of these labels.

### 4.1.3. Assessment of the main experimental paradigm (word-spotting)

#### 4.1.3.1. Strengths

Word-spotting offers ecological validity, in that it closely resembles the natural task of spotting words in continuous input without knowing in advance what those words are. Also, as discussed in the literature review, the word-spotting paradigm has the capacity to highlight phonotactic effects on online speech segmentation.

---

[17]CEGEP: Collège d'enseignement général et professionnel (College of General and Vocational Education).

[18]'Advanced' participants further included 2 students with no prior immersion in Anglophone settings. Since priority was given to the quantitative/experimental assessement of L2 proficiency, these 2 participants had to labelled 'Advanced' in light of their good experimental results.

[19]These 2 undergraduate students turn out to be the 2 additional 'Advanced' students mentioned in the above footnote. One was a Montréal-based, CEGEP student; the other was an undergraduate student enrolled at the University of Alberta in Francophone courses only. In both cases, the 'Advanced' status was confirmed by the fact that their experimental scores surpassed those of one of the MBA students.

4.1.3.2. Artifacts and solutions

Word-spotting is also difficult. Error rates tend to be high, and non-native speakers may not be able to spot words that they do not know (McQueen, 1998).

Two methods, in this study, helped prevent such floor effects. First, the easier task of lexical decision (Experiment 3) offered reliable evidence of how sensitive the learners might be to lexical and phonotactic cues. Second, participants were warned that, in word-spotting, each target word would start with a consonant, would be one syllable long, and would appear at the end of a stimulus string. These hints, with respect to stimuli like 'coafpage', prevented unwanted responses like 'coaf' (a nonword) or 'age' (a non target word).

Another artifact of word-spotting is that it may tap mechanisms of compound processing rather than multiword/utterance segmentation (Libben, p.c.). Thus, if listeners manage to spot the target word 'code' in 'thoafcode', it may be because they hear 'thoafcode' as a sort of 'code' (by identifying the head of a fictional compound) and not because they segment the utterance 'thoaf code' via lexical or phonotactic cues (e.g., the HF word 'code' and the CC3 onset [fk]). This challenge can be addressed in two ways.

First, evidence of contrastive response patterns under the Aligned vs. Misaligned conditions (e.g., 'thoafcode' vs. 'thoascode') would suggest that the word-spotting task had tapped the speech segmentation mechanisms posited in this research.

Second, a confound between compound processing and speech segmentation would seem to involve the morphological size effects typical of compound processing. Target words were therefore tagged for:

1) compound family size, referring to the number of bisyllabic compounds that contain a target word as head or modifier (e.g., 'barcode', 'code-key');

2) compound positional family size, referring to the number of bisyllabic compounds that contain the target word as head only (e.g., 'barcode').

A two-step rationale ensued:

1) should neither of these parameters display a significant effect on the reported word-spotting data, then one could reject the conclusion that word-spotting had tapped compound processing rather than speech segmentation mechanisms;

1) should neither of these parameters display a significant effect on the reported word-spotting data, then one could reject the conclusion that word-spotting had tapped compound processing rather than speech segmentation mechanisms;

2) should either of these parameters show an effect, then evidence for an interaction of lexical and phonotactic cues could still yield interesting conclusions about multiword segmentation in the auditory mode.

### 4.1.4. Language modes

The testing of bilinguals can be challenged by the degree to which the participants' languages are activated during testing (Grosjean, 2001). Bilinguals can thus be said to be in monolingual mode (i.e., with only one of their two languages being activated), in bilingual mode (with both of their languages being activated), or in intermediate mode.

The monolingual mode is preferrable in studies (such as ours) that focus on L2 psycholinguistic processes–as pointed out by Marian and Spivey (2003). However the monolingual mode seems difficult to achieve, as a bilingual's L1 may never be fully deactivated (Grosjean, 2001: 7). In the present case, a strict monolingual mode was in fact not entirely desirable since the study was interested in potential transfer effects from the L1. Efforts were made, nonetheless, to ensure that participants were as close as possible to a monolingual mode. Thus, following Grosjean's indication (2001: 5) that a bilingual's position on the language mode continuum can be influenced by his/her knowledge of the interlocutor's language background, participants communicated with the experimenter in L2 English only, with no possibility to switch back to L1 French.

### 4.1.5. Statistical testing

In the following Results sections, statistical reports were interpreted on the primary basis of by-subjects analyses of variance (ANOVAs).

On this basis, if a main or interaction effect involving more than two stimulus conditions was significant for the participants only or for both participants and items, these conditions were compared in the form of by-subjects t-tests. If a main or interaction effect involving more than two stimulus conditions was significant for items only, these conditions were compared in the form of by-items t-tests. Unless reported otherwise, these additional t-tests were planned rather than post-hoc tests.

A conservative approach was taken to the investigation of all these condition-pair comparisons, using Bonferroni adjustments rather than uncorrected t-tests.

Note also that, if in the context of a reported ANOVA higher-order interactions are not discussed, it is because they had been found to be statistically non-significant.

## 4.2. EXPERIMENT 3: LEXICAL DECISION

### 4.2.1. Rationale

Lexical decision accuracy scores contributed to the assessment of L2 proficiency.

The main aim of this experiment, however, was to prepare the word-spotting experiment (Experiment 4) by checking:

1) the validity of the stimuli tested by the variables Frequency, Cognacy and Onset Type;
2) the plausibility that Francophone ESL learners would be responsive to the potential lexical cues (Frequency and Cognacy) and phonotactic cues (Onset Type) of interest;
3) the plausibility of L1 French transfer on L2 English word recognition. This last check involved a pilot run of the lexical decision task with a control group of Chinese ESL learners, Chinese being dramatically more different from Chinese than from French.

With regard to Frequency and Cognacy, recall the conditions HF vs. LF and YESCOG vs. NOCOG. As for Onset Type, recall the conditions C (as in 'k̲ite'), CC1 (unmarked in English and French, as in 'black' or 'proun'), CC2 (marked in French but not in English, as in 's̲k̲in' or 's̲p̲ile'), CC3 (illegal in English and French, as in 't̲m̲oul').

### 4.2.2. Materials

#### 4.2.2.1. Words

*Simple-onset words.* Ninety-six simple-onset (or C) words like 'kite' were equally split into words starting in [k, l, m, n, p, r], HF vs. LF and YESCOG vs. NOCOG words.

Frequency was operationalized under the premise that L2 and L1 usage frequencies were proportionate. Simple-onset words were thus drawn from Celex so that their frequency values maximized the HF vs. LF contrasts (Table 12).

Table 12. Frequency values of Low vs. High Frequency simple-onset words

| | Low Frequency (LF) Words | High Frequency (HF) Words |
|---|---|---|
| Mean | 2 | 397 |
| Standard Deviation | 2 | 620 |
| Range | 0-9 | 16-3783 |

Note. Frequency values originally reported by Celex per 1.3 million words of spoken text

Simple-onset words were labelled YESCOG if they appeared in the Robert-Collins dictionary with similar forms and similar meanings in English and French (e.g., 'North', for French 'Nord') and if they appeared in Hammer and Giauque's (1989) English-French cognate dictionary. The Petit Robert further helped to validate 8 words from 'franglais', the rapidly expanding stock of English loanwords in contemporary French[20]. Treating loanwords as cognates may seem unorthodox, but was made necessary by the need for target words specifiable in terms of Frequency (high or low) and Onset Type (C or, as discussed next, CC1/CC2) as well as in terms of Cognacy. A flexible interpretation of the notion of cognacy seemed further justified by evidence for graded cognacy ratings by bilinguals (de Groot & Nas, 1991; Friel & Kennison, 2001).

Celex data showed that HF and LF simple-onset words were matched for orthographic neighborhood [$t(92) = 1.08$, $p = 0.28$], phonological neighborhood [$t(92) = 1.08$, $p = 0.28$][21], word length in letters [$t(94) = 1.47$, $p = 0.15$], word length in phonemes [$t(94) = 0.75$, $p = 0.46$]. MRC data showed that simple-onset LF words had significantly higher concreteness and imageability ratings [$t(63) = 3.49$, $p = 0.001$; $t(66) = 3.61$, $p = 0.001$] than simple-onset HF words. This could be because the MRC lacked concreteness data for 25 LF words vs. 6 HF words and imageability data for 24 LF words vs. 4 HF words. In any case, higher concreteness and imageability for LF words seemed like positive factors, as they might help make such words more accessible to L2 learners.

[20]Here are these words, along with bracketed examples of relevant Google entries: 'cash' ("forme de cash"), 'loft' ("grand loft"), 'look' ("le look de Montréal"), 'man' ("one-man show"), 'must' ("un must à Québec"), 'news' ("les news de 20 heures"), 'night' ("Montréal by night"), 'made' ("créativité made in Québec").

[21]These neighborhood tests involved less degrees of freedom as Celex lacked values for 'made' and 'lap'.

*Complex-onset words.* Forty-eight complex-onset words (e.g., 'plead') were equally split into words in onsets [bl, br, gl, gr, kl, kr, pl, pr, sk, sm, sn, sp], HF vs. LF and YESCOG vs. NOCOG words. Frequency and Cognacy were implemented as above.

HF and LF complex-onset words were matched for orthographic neighborhood [$t(46) = 0.33$, $p = 0.75$], phonological neighborhood [$t(46) = 0.85$, $p = 0.40$], word length in letters [$t(46) = 0.57$, $p = 0.57$] and in phonemes [$t(46) = 1.00$, $p = 0.32$]. They were also balanced for concreteness [$t(30) = 0.02$, $p = 0.99$] and imageability [$t(32) = 0.35$, $p = 0.73$], though these last tests should be viewed with caution as a number of concreteness and imageability data were missing from the MRC database.

4.2.2.2. Nonwords

*Simple-onset nonwords.* Sixty C nonwords like 'kaygue' [kejg] were taken from the ARC Nonword Database (Rastle, Harrington & Coltheart, 2002) if they resembled hence potentially competed with C words (e.g., 'case'). They thus contained 6 of the 8 nucleus vowels present in the C words (i.e., 2 monophthongs and 4 diphthongs). These nonwords were also equally split between the onset consonants [k, l, m, n, p, r] and matched for phonological neighborhood across these consonants [$F(5) = 1.08$, $p = 0.38$].

*Complex-onset nonwords.* Eight complex-onset nonwords (24 CC1 items like 'klug', 24 CC2 items like 'smag', 32 CC3 items like 'fkev') were drawn from the ARC Database. They closely resembled C nonwords, in that they showed the same proportions of nucleus vowels as C nonwords (e.g., 10% for the vowel [eɪ], as in 'clayne' and 'kaygue').

Onset type significantly affected phonological neighborhood [$F(2,77) = 160.84$, $p = 0.0001$]. Post-hoc, Bonferroni-adjusted t-tests (per-test alpha level set at 0.016) showed that while CC1 and CC2 nonwords did not significantly different (i.e., were balanced) for phonological neighborhood [$t(46) = 1.95$, $p = 0.06$], CC3 nonwords had significantly less phonological neighbors than CC1 nonwords [$t(25) = -15.91$, $p = 0.0001$] and CC2 nonwords [$t(25) = -13.63$, $p = 0.0001$]. No wonder since CC3 nonwords, unlike CC1 and CC2 nonwords, included onset clusters not found in English words.

4.2.2.3. Final stimulus set

The set comprised 144 words (96 with simple onsets, 48 with complex onsets) and 140 nonwords (60 with simple onsets, 80 with complex onsets). These proportions not only allowed the testing of lexical and phonotactic skills. By mixing nonwords of various syllabic formats, they also prevented floor effects, which may have arisen as the task used no priming paradigm and only used morphologically simple target words.

4.2.2.4. Recordings

The stimuli were recorded in a sound-insulated booth, via the Alesis Masterlink ML-9600 High Resolution Master Disk Recorder, by reading a list of citation forms. The speaker, a female speaker of Canadian English (Saskatchewan), was an honours linguistics student, and so her rendition of odd-sounding items like 'tlar' was easily rehearsed using phonological terminology (e.g., "please try this onset cluster again").

### 4.2.3. Procedure

Participants were tested one by one. A portable iMac computer running PsyScope 1.2.2 (Cohen, MacWhinney, Flatt, & Provost, 1993) was used to randomize stimuli, play the stimuli, and collect reaction times with millisecond accuracy. Each trial proceeded in the following three steps: 1) an asterisk appeared in the center of the computer screen for 500 ms ; 2) the asterisk was then removed; 3) one of the randomly selected stimuli was presented over a pair of Sony MDR-V600 Dynamic Stereo Headphones. Reaction times were measured from the onset of the stimulus to the onset of the button press response.

Prior to the experiment, participants read the instructions from the computer screen. The instructions were in English in order to inforce the English mode discussed earlier. Participants were instructed to respond as quickly and as accurately as possible. If they thought that the item was a real English word, they were to press the green button labeled 'Yes'. If they thought otherwise (i.e., that the item was a nonword), they were to press the button labeled 'No'. Once presented with these instructions, each participant received 10 practice trials, which were not included in the final analysis.

### 4.2.4. Results and discussion

4.2.4.1. Data processing

Accuracy scores were standardly computed as 0 for 'no' presses in response to real words or 'yes' responses to nonwords, and as 1 in the alternate cases.

RTs were kept for accurately recognized words and for all nonwords, as both accurate and erroneous nonword responses revealed how ESL learners treated English-like items of various phonotactic statuses. Outlier word and nonword RT data (i.e., below 300 ms or above 2000 ms) were removed as they probably reflected variables other than those of interest (e.g., guesses, fatigue). Also excluded were values more than 2 standard deviations above or below the mean across conditions and participants. Among the original RTs, these values represented respectively 9% and 13% of the data for the Chinese participants (pilot task), 5% and 6% for the French participants (main task).

4.2.4.2. Pilot results of the Chinese participants

*Frequency*. Figure 1 (next page) shows that, in comparison to LF words, HF words standardly got higher accuracy scores [$F1(1,25) = 35.91$, $p = 0.0001$; $F2(1,38) = 43.96$, $p = 0.0001$] and lower RTs [$F1(1,25) = 35.24$, $p = 0.0001$; $F2(1,31) = 15.21$, $p = 0.0001$].

*Cognacy*. French-English cognacy unsurprisingly had no significant effects on Chinese learners, neither in connection with accuracy scores [$F1(1,25) = 0.40$, $p = 0.53$; $F2(1,32) = 0.94$, $p = 0.34$] or with RTs [$F1(1,25) = 0.03$, $p = 0.85$; $F2(1,31) = 0.32$, $p = 0.57$].

*Onset Type*. For words, Onset Type did not significantly affect accuracy scores [$F1(2,50) = 2.23$, $p = 0.12$; $F2(2,13) = 0.65$, $p = 0.54$] nor RTs [$F1(2,50) = 0.10$, $p = 0.90$; $F2(2,10) = 1.21$, $p = 0.34$].

For nonwords, Onset Type had significant by-subjects and by-items effects on RTs [$F1(3,75) = 3.87$, $p = 0.01$; $F2(3,77) = 2.79$, $p = 0.05$] and significant by-subjects effects on accuracy scores [$F1(3,75) = 7.41$, $p = 0.0001$ ; $F2(3,77) = 1.72$, $p = 0.17$].

Bonferroni-adjusted by-subjects t-tests (per-test alpha of 0.025) specified that significant contrasts of accuracy arose between C vs. CC3 nonwords [t(25) = -3.12, p = 0.004] and between CC1 vs. CC3 nonwords [t(25) = -3.83, p = 0.001]. Also, Bonferroni-adjusted by-items t-tests (per-test alpha of 0.025) showed significant RT differences, if not between C vs. CC3 nonwords [t(31) = 2.04, p = 0.05][22], at least between CC1 vs. CC3 nonwords [t(23) = 3.40, p = 0.002]. Overall, these results suggest that CC3 nonwords were processed most accurately and fastest. This conclusion, explainable by the fact that CC3 onsets had been designed to be the most infelicitous hence most salient onsets, validates the way that phonotactic variations were implemented in the stimulus set.

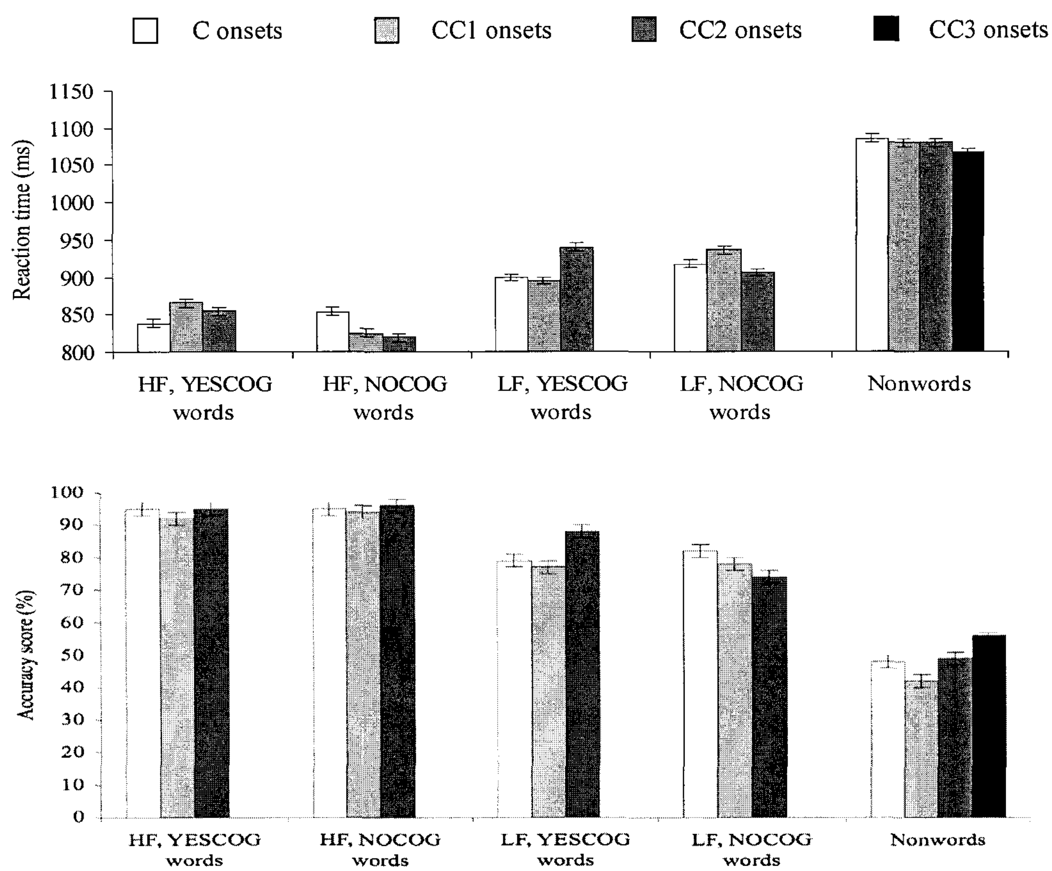



Figure 1. Pilot lexical decision results of the Chinese participants

[22]This is the only case in the study where the conservative, Bonferroni adjustment of a planned t-test (as announced in Section 4.1.5) ended up cancelling out an otherwise significant result. By fear of an ensuing Type-II error, the author wishes to emphasize that the Chinese learners did respond to CC3 nonwords faster than to C nonwords, in the continuity of the general response pattern described in this paragraph.

4.2.4.3. Results of the Francophone participants: Effects of stimulus variables

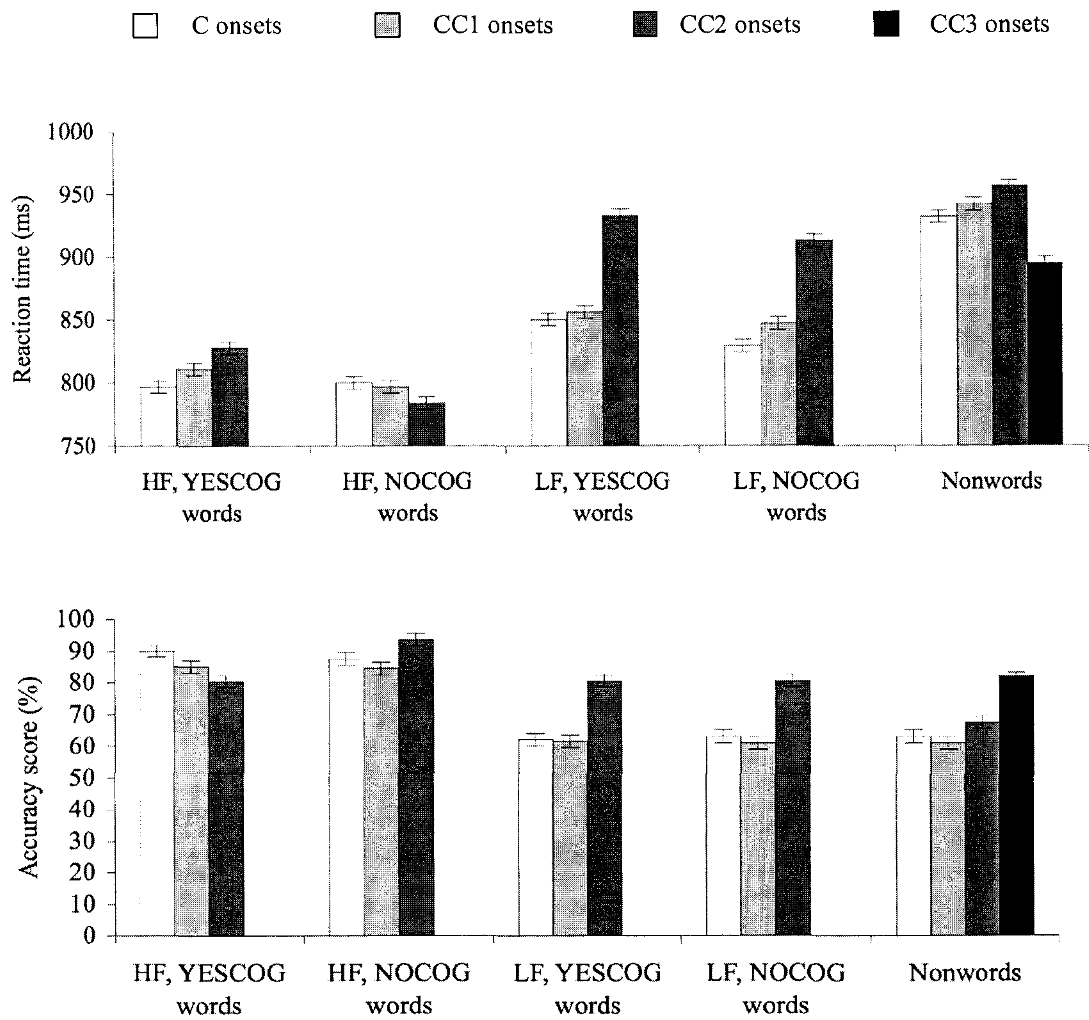



Figure 2. Lexical decision results of the Francophone participants (by stimulus variable)

*Frequency*. Figure 2 shows that, in comparison to LF words, HF words standardly got higher accuracy scores [F1(1,17) = 63.36, p = 0.0001; F2(1,93) = 64.60, p = 0.0001] and lower RTs [F1(1,17) = 73.30, p = 0.0001; F2(1,140) = 31.49, p = 0.0001].

*Cognacy*. Cognacy had no significant effects on accuracy scores [F1(1,17) = 0.002, p = 0.96; F2(1,78) = 0.50, p = 0.48] but significant by-subject effects on RTs [F1(1,17) = 6.44, p = 0.02; F2(1,164) = 0.40, p = 0.53]. A closer look at Cognacy data shows indeed a slight non-cognate advantage for NOCOG words over YESCOG words (842 ms vs. 860 ms). Cognacy may thus have acted, though moderately, as a source of nuisance. This is in line with a language non-selective take on bilingual processing, whereby the processing of cognates is hindered by the association of these words with L1 words.

The Cognacy x Frequency interaction did not significantly affect RTs [$F1(1,17) = 0.19$, $p = 0.67$; $F2(1,185) = 0.05$, $p = 0.83$] but did significantly affect accuracy scores on a by-subject basis [$F1(1,17) = 17.57$, $p = 0.001$; $F2(1,185) = 2.11$, $p = 0.15$]. Bonferroni-adjusted, by-subject t-tests (per test alpha of 0.012) specified that: 1) given a set level of Cognacy, Frequency significantly affected accuracy scores [YesCog-HF vs. YesCog-LF: $t(35) = 6.88$, $p = 0.0001$; NoCog-HF vs. NoCog-LF: $t(35) = 9.09$, $p = 0.0001$]; 2) Cognacy had a significantly negative effect under the HF condition [HF-YesCog vs. HF-NoCog: $t(35) = -2.58$, $p = 0.01$] and no significant effect under the LF condition [LF-YesCog vs. LF-NoCog: $t(35) = 1.56$, $p = 0.13$].

These analyses together point out that English-French Cognacy did not facilitate the lexical decision performances of the Francophone ESL learners.

*Onset Type (words)*. Onset Type had no significant effects on scores [$F1(2,34) = 1.07$, $p = 0.35$; $F2(2,10) = 0.47$, $p = 0.64$] but had significant effects on RTs by-subjects [$F1(2,34) = 12.86$, $p = 0.0001$; $F2(2,10) = 3.61$, $p = 0.07$]. Bonferroni-adjusted, by-subject t-tests (per-test alpha of 0.017) showed that RTs did not significantly differ between C and CC1 words [$t(35) = 0.89$, $p = 0.38$] but were significantly higher for CC2 words than C words [$t(35) = 4.63$, $p = 0.0001$] and CC1 words [$t(35) = 3.07$, $p = 0.004$].

The Onset Type x Frequency interaction did not significantly affect accuracy scores [$F1(2,34) = 0.54$, $p = 0.59$; $F2(2,185) = 0.40$, $p = 0.67$] but did significantly affect RTs [$F1(2,34) = 12.60$, $p = 0.0001$; $F2(2,185) = 3.68$, $p = 0.03$]. Two of 3 Bonferroni-adjusted, by-subject t-tests (per test alpha of 0.017) showed that LF-CC2 words had been responded to significantly more slowly than LF-CC1 words [$t(33) = 3.43$, $p = 0.002$] and than LF-C words [$t(33) = 4.39$, $p = 0.0001$].

The Onset Type x Frequency x Cognacy interaction, while not significantly affecting RTs [$F1(2,34) = 2.19$, $p = 0.13$; $F2(2,185) = 0.27$, $p = 0.76$], significantly affected by-subject accuracy scores [$F1(2,34) = 15.10$, $p = 0.0001$; $F2(2,185) = 2.92$, $p = 0.06$]. One of 3 Bonferroni-adjusted by-subjects t-tests (per test alpha of 0.017) showed that YesCog-HF words had been responded to significantly less accurately if they started with CC2 onsets than with C onsets [$t(35) = -3.31$, $p = 0.002$].

These analyses together indicate that CC2 words were hardest to recognize.

*Onset Type (nonwords)*. Onset Type had significant effects on nonword accuracy scores [$F1(3,51) = 27.59$, $p = 0.0001$; $F2(3,77) = 5.99$, $p = 0.001$] and RTs [$F1(3,51) = 18.39$, $p = 0.0001$; $F2(3,77) = 13.67$, $p = 0.0001$]. Table 13 shows the outcome of all the post-hoc, Bonferroni-adjusted by-subjects t-tests (per test alpha of 0.008 for both accuracy score tests and RT tests). In light of this table, two performance rankings emerge for nonwords: the accuracy ranking CC1 < C < CC2 < CC3 and the speed ranking CC2, CC2, CC1 < C < CC3. The slight difference between these rankings is accountable for in terms of a speed-accuracy tradeoff whereby CC2 nonwords were simpler to reject than nonwords starting with other types of legal onsets, so at the risk of slower processing.

Table 13. Onset Type effects on nonword processing in the main lexical decision task

| Accuracy scores | Reaction times |
|---|---|
| C (63%) vs. CC1 (61%):<br>$t(35) = 2.32$, $p = 0.03$ | C (935 ms) vs. CC1 (945 ms)<br>$t(35) = -1.24$, $p = 0.22$ |
| C (63%) vs. CC2 (69%)<br>$t(35) = -3.00$, $p = 0.005$ | C (935 ms) vs. CC2 (960 ms)<br>$t(35) = -3.21$, $p = 0.003$ |
| C (63%) vs. CC3 (82%)<br>$t(35) = -7.30$, $p = 0.0001$ | C (935 ms) vs. CC3 (898 ms)<br>$t(35) = 3.24$, $p = 0.003$ |
| CC1 (61%) vs. CC2 (69%)<br>$t(35) = -4.18$, $p = 0.0001$ | CC1 (945 ms) vs. CC2 (960 ms)<br>$t(35) = -1.55$, $p = 0.13$ |
| CC1 (61%) vs. CC3 (82%)<br>$t(35) = -8.04$, $p = 0.0001$ | CC1 (945 ms) vs. CC3 (898 ms)<br>$t(35) = 4.41$, $p = 0.0001$ |
| CC2 (69%) vs. CC3 (82%)<br>$t(35) = -4.59$, $p = 0.0001$ | CC2 (960 ms) vs. CC3 (898 ms)<br>$t(35) = 6.18$, $p = 0.001$ |

Note. Recall the categories C (simple onsets), CC1 (unmarked in English & French), CC2 (marked in French, unmarked in English), CC3 (illegal in English & French).

4.2.4.4. Results of the Francophone participants: Effects of Proficiency

Figure 3 shows the effects of Proficiency along with the stimulus variables. Accuracy scores are ignored as they served to build the Proficiency groups (see Section 4.1.2.1).

*Proficiency (words)*. Advanced learners were faster than Intermediate learners though only significantly so by items [$F1(1,17) = 1.46$, $p = 0.24$; $F2(1,185) = 7.20$, $p = 0.008$].

The Proficiency x Cognacy interaction was not significant for word RTs [$F1(1,178) = 0.002$, $p = 0.97$; $F2(1,185) = 0.31$, $p = 0.58$]. Nor was the Proficiency x Frequency interaction [$F1(1,178) = 0.10$, $p = 0.75$; $F2(1,185) = 0.002$, $p = 0.97$].

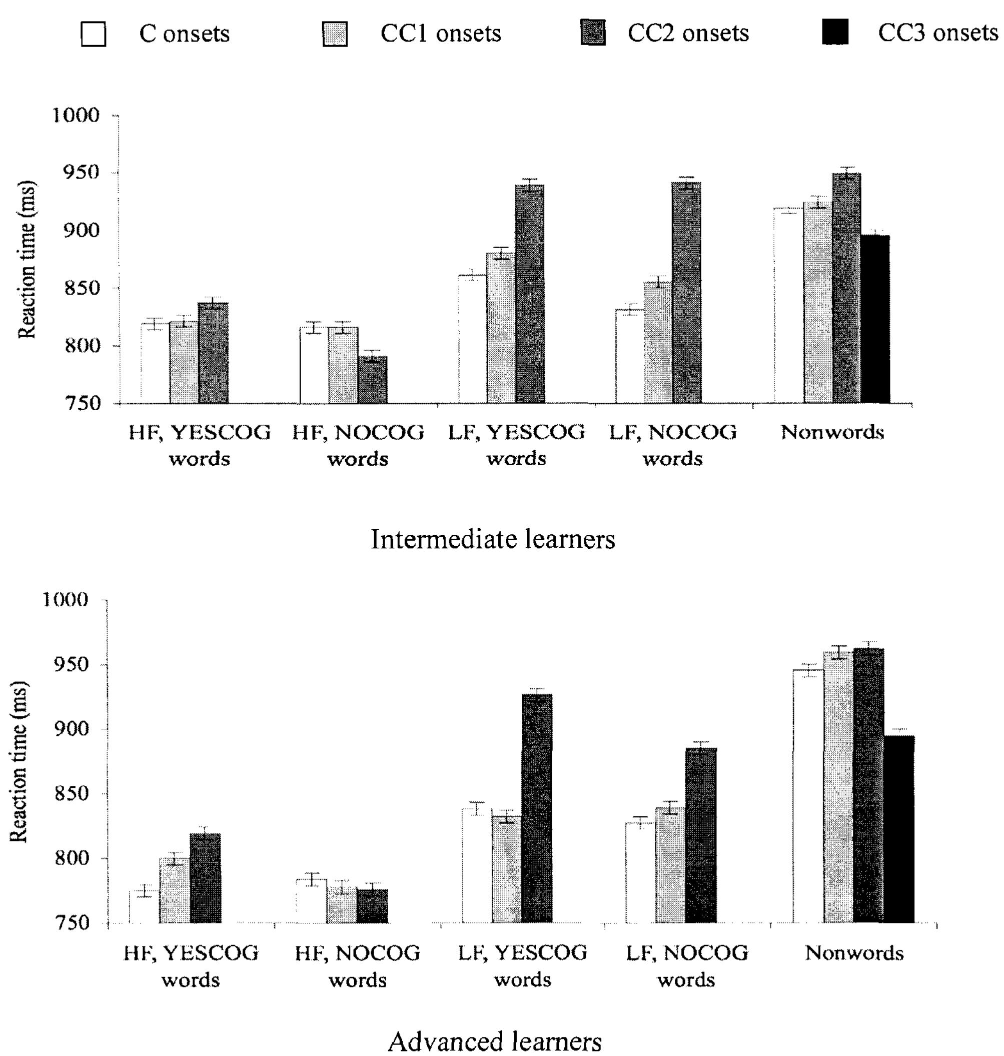



Figure 3. Lexical decision RTs of the Francophone participants (by Proficiency group)

*Proficiency (nonwords)*. Advanced learners were faster than Intermediate learners though only significantly so by items [$F1(1,17) = 0.34$, $p = 0.57$; $F2(1,136) = 10.79$, $p = 0.001$].

The Proficiency x Onset Type interaction was not significant for nonword RTs [$F1(3,51) = 1.91$, $p = 0.14$; $F2(3,136) = 1.78$, $p = 0.15$].

Sonority effects on CC3 responses were analyzed for Advanced vs. Intermediate learners. This analysis reused the sonority scale 'Plosives–Fricatives/Affricates–Nasals–Liquids–Glides–Vowels', whereby segments located further to the left of the scale are preferrable onsets and those closer to the right are preferrable nuclei. Given the cross-linguistic significance of this scale (see, e.g., Parker, 2002), and given that French specificities concerning sonority rankings seem minimal[23], what was at stake in this analysis was whether L2 listeners might exploit presumably universal rather than L1- or L2-specific phonotactic cues in L2 processing. Three types of CC3 onsets were thus examined, with reference to the principles of Sonority Sequencing (SS) ("sonority must rise from syllable onset to syllable peak") and Minimal Sonority Distance (MSD) ("the members of a CC cluster must be of at least two degrees apart on the Sonority Scale"):

1) 'tolerable' CC onsets, violating neither the SS or MSD principle (viz., [pn, sr, tl, tm]);
2) 'bad' CC onsets, violating the MSD principle (viz., [kf, ps]);
3) 'very bad' CC onsets, violating both SS and MSD principles (viz., [fk, fp, ln, mz, pt]).

Learners were overall slower with 'very bad' onsets than with 'bad' onsets (see left halves of the graphs in Figure 4). This difference was larger for Advanced learners (14% and 14 ms) than for Intermediate learners (5% and 10 ms), though the Proficiency x CC3 Subtype interaction was not significant[24]. This moderate loss of speed from 'bad' to 'very bad' onsets has two possible origins: attention and/or guessing. Attention to sonority variations might have played a role in the task, despite the time pressure put on participants; specifically, 'very bad' onsets may have called for more attention. But guessing strategies could also have underlied the processing of CC3 nonwords, as suggested by the plunge of RTs in the Intermediates' responses to 'tolerable' onsets.

[23]See, for instance, Rialland (1994), for the proposal that French glides can either be onsets or nuclei.

[24]In a two-way ANOVA (Proficiency x CC3 Subtype), with RT as dependent variable, CC3 Subtype had significant main effects on a by-subjects basis [$F1(2,34) = 3.66$, $p = 0.04$; $F2(2,18) = 2.52$, $p = 0.11$], but no significant interaction with Proficiency [RT: $F1(2,34) = 1.71$, $p = 0.20$; $F2(2,18) = 0.30$, $p = 0.47$].

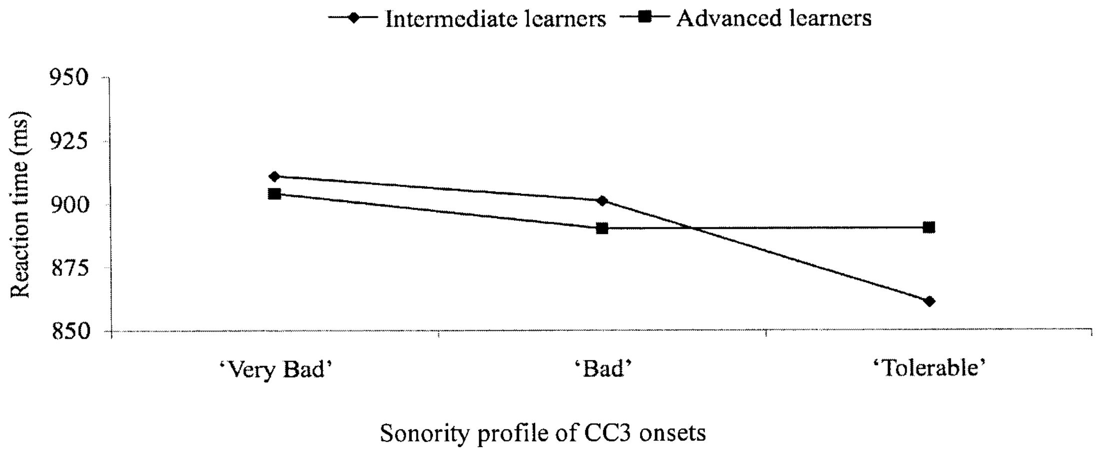



Figure 4. Effects of sonority on learner performance in the main lexical decision task

4.2.4.5. Summary and conclusions

These variable-by-variable findings point out the types of cues that may help the auditory recognition of English words by Francophone ESL learners.

With respect to potential lexical cues: while Cognacy fails to help (and even slows down) performance, Frequency proves significantly helpful, though there was no evidence that Frequency increases facilitated one Proficiency group more than the other.

With respect to phonotactic cues, it remains unclear whether learners processed illegal onsets by attending sonority variations or by guessing. There was also no evidence that Intermediate and Advanced learners experienced Onset Type variations differently. At the same time, this experiment confirms that L2 spoken word recognition is sensitive to the comparative phonotactics of an L1 and and L2, as shown by evidence that: 1) CC2 words were hardest to recognize; 2) CC2 nonwords, though not as easy to reject as CC3 nonwords, were more accurately rejected than nonwords starting with other types of legal onsets.

The next experiment will indicate whether these conclusions, based on isolated words, still apply when these words are presented within larger, segmentable strings.

## 4.3. EXPERIMENT 4. WORD-SPOTTING

### 4.3.1. Rationale

Word spotting accuracy scores were needed to assess L2 proficiency. But the task mainly sought to test learner sensitivity to Frequency, Cognacy and Onset Type during online L2 speech segmentation. This task was thus the central experiment of the study.

### 4.3.2. Materials

#### 4.3.2.1. Carrier strings

*General format.* The carrier strings (or carriers) were nonsense CVCCVC(C) strings like 'voofcase' or 'vooscase', ending in target words reused from Experiment 3 (e.g., 'case').

*Design procedure.* The first design step consisted in building string-initial CVC syllables like 'voof' or 'voos'. For this purpose, the vowels [ɔ, u, aw, ow, aj, ej] were used as CVC nuclei, and appeared in this position in approximately equal proportions across all carrier groups (e.g., 'HF-YESCOG Aligned' or 'LF-NOCOG Misaligned'). While reducing the risk of French-sounding stimuli[25], these tense vowels could act as legal CV nuclei, thus giving listeners a chance to make albeit erroneous segmentations like 'voo-scase'. Also, as in McQueen (1998), these vowels were positioned so as to create a sense of (quasi-) homophony in Aligned-Misaligned pairs like '<u>voof</u>case'-'<u>voos</u>case', while at the same time creating a certain phonetic diversity hence reduced monotony across the set (see, e.g., the string-initial CVCs 'fawp', 'thea' or 'lou'). The next design step consisted in adding to these CVCs the legal onsets [bl, fl, gl, kl, pl, br, fr, gr, kr, pr, sk, sm, sn, sp] or illegal onsets [fk, fp, km, pn, sr, tl], as well as the CVC target words (e.g., 'case').

[25]Out of the same concern, [i] was not used because of its resemblance with French [i], a resemblance possibly more acute than the one between English and French [ɔ, u]. Also, the diphthong [oj] was avoided because it is the least frequent diphthong in English (Fry, 1947; cited in Crystal, 1995: 125) and thus risked becoming overtly salient, which could have introduced a bias for CV.CCVC or CVC.CVC responses.

In total, 96 pairs of Aligned vs. Misaligned carriers were created (Appendix H).

*Control measures.* Three measures sought to minimize the interference of extraneous lexical cues on the processing of the carriers. First, CVC nonwords from Experiment 3 (e.g., 'poun' or 'brett') were not reused lest they risked gradually acquiring a real-word status from the point of view of an L2 learner. Second, a check was made that the vast majority of string-initial CVCs (e.g., 'louk', [lawk]) and embedded CV portions (e.g., 'lou', [law]) appeared in the ARC Nonword Database. Third, only a small portion of (C)CV(C) non-target words were tolerated inside the carrier strings, providing that they were simultaneously listed as obsolete/dialectal in the Oxford English Dictionary and absent from the American Heritage Dictionary (e.g., 'taise', a measure unit) or that they belonged to a highly specialized vocabulary presumably outside the reach of university-level ESL learners (e.g., 'skeg', a nautical term).

4.3.2.2. Filler and practice strings

Fifty filler strings containing no real words were built using the same procedure as for the carrier strings. Half these fillers were based on illegal onsets like [tl], thus qualifying as Aligned strings (e.g., 'fowt.leck'); the other half were based on legal onsets like [bl], thus qualifying as Misaligned strings (e.g., 'vai.blan'). In order to prevent listeners from using deliberate segmentation strategies through the task (e.g., interpreting the recurrence of illegal onset [tl] as a signal of upcoming target words), the legal and illegal onsets of interest appeared in the fillers in the same proportions as in the carriers.

Ten practice strings were finally designed, equally split between Aligned and Misaligned strings, and made of target words other than the experimental targets.

4.3.2.3. Recordings

All materials were recorded in a soundproof booth by the Canadian English speaker recruited for Experiment 3. Prior rehearsal of all materials, in the company of the experimenter, allowed to ensure that the stimuli displayed the strong-weak prosodic pattern typical of bisyllabic English phrases and compounds. Maintaining this prosodic pattern across the set sought to reduce the risk of prosodic cues interfering with the lexical vs. phonotactic cues under study.

4.3.2.4. Stimulus lists

For the main, non-practice trials, two stimulus lists were made to counterbalance the Alignment and Misalignment conditions for the carrier strings. Thus, while Misaligned 'lou.frich' appeared in List 1, Aligned 'dous.rich' appeared in List 2.

Each list contained 96 carriers and 48 fillers. Three factors motivated this lesser proportion of fillers. First, pretesting with fellow experimenters from the Centre for Comparative Psycholinguistics (University of Alberta) suggested that 48 fillers would suffice to conceal the phonotactic cues of interest from L2 leaners. Second, stimulus lists had to be reduced to a strict minimum because of the financial and logistic need to spend 1 hour maximum with each participant[26]. Third, it remained possible to pseudorandomize the 96 carriers and 48 fillers in such a way that two Aligned or Misaligned carriers never appeared in direct succession, and so that no more than three carriers or fillers ever appeared successively (e.g., Dumay, Frauenfelder & Content, 2002).

### 4.3.3. Procedure

Much of the procedure from Experiment 3 was kept: participants were tested one by one; an iMac running PsyScope 1.2.2 presented stimuli and collected RTs; each trial consisted in a 500 ms warning star followed by a headphone presentation of the stimulus; and RTs were measured from the onset of the stimulus to the onset of the button press.

English instructions on the iMac screen asked participants to perform either one of two pairs of actions: 1) press the green button as fast as possible if they thought that a stimulus string contained a real English word, then name that word aloud; 2) press the red button as fast as possible if they thought that a stimulus string did not contain a real English word, and confirm this response by saying 'no word'. The first of these pairs of

[26]This hour was split between an introduction to the proposed research, the background questionnaire, and Experiments 3 to 6.

actions is standard word-spotting procedure. The second pair of actions was added for the purpose of testing the learners' ability to reject items devoid of real words (as part of the Proficiency assessment set off in Experiment 3).

Participants were told in advance that real words, if present, would appear string-finally and would start with consonants. These two hints, as already discussed in Section 4.1.5, sought to reduce the risk of floor effects in an already challenging task.

Prior to the experiment, participants were given 10 practice trials, which were not included in the final analysis. These practice trials were repeated if needed.

### 4.3.4. Results and discussion

#### 4.3.4.1. Data processing

*Scoring procedures.* Two scoring procedures were used: 1) 'proficiency scoring', which served to refine the assessment of L2 lexical proficiency; 2) 'WS (word-spotting) scoring', which assessed word-spotting performance as such.

Proficiency scoring was more liberal than WS scoring. Given the difficulties inherent to word-spotting, it seemed expectable that learners might confuse intended target words with other English words. Moreover, the detection of such unintended English words seemed a valuable indicator of the subject's L2 lexical knowledge.

For WS scoring, all RTs were recorded, whether the red or green button had been pressed. A score of 0 was assigned in any one of these contexts: 1) if the corresponding RT fell outside the 200-2500 ms range (and was then discarded from later analysis[27]); 2) if the red button was erroneously pressed; 3) if the green button was appropriately pressed but no oral response to a carrier string was given; 4) if the green button was appropriately pressed but a word other than the intended target word was named. A score of 1 was assigned if and only if a participant accurately named the string-final target.

[27]The outlier margin of 2500 ms, while reflecting the fact that the task had tapped online word processing under time pressure, was preferred to the 2000 ms of the lexical decision task because it better acknowledged the challenges specific to word spotting in a second language.

For Proficiency scoring, green and red button presses were equally considered. A score of 1 was given if the word named after a green button press was the target, or if the red button was accurately pressed (i.e., for strings devoid of real words). A score of 0 was given if an accurate green button press came with no naming response. A score of 0.5 was given if the named word differed by one phoneme from the target (e.g., 'ran' instead of 'run') or if the subject changed a string-final nonword (e.g., 'poot', in 'zayspoot') into a real word (e.g., 'boot'). The idea behind these 0.5 scores was to show in the Proficiency assessment when learners appeared to know real English words.

#### 4.3.4.2. Results of the Francophone participants: Effects of stimulus variables

*Preliminaries.* The following discussion is strictly concerned with error rates and RTs gathered through the standard WS scoring procedure.

An initial ANOVA with the variable Test List (along with Onset Type, Cognacy, Frequency and Proficiency) found no main or interaction effects for this variable (all Fs $< 1.40$, $p. > 0.24$). RT and error rate data were therefore collapsed across lists in subsequent analyses, and so the per-variable tests reported below in connection with Figure 5 relate to the ANOVA Onset Type x Cognacy x Frequency x Proficiency.

Figure 5 identifies two standard word-spotting patterns:

1) a processing advantage for target HF words over target LF words;

2) a processing advantage (i.e., overall lower error rates and RTs) for Aligned strings (built on CC3 onsets) over Misaligned strings (built on CC1 or CC2 onsets).

*Frequency.* Frequency significantly affected error rates [$F1(2,17) = 227.85$, $p = 0.0001$; $F2(1,54) = 101.80$, $p = 0.0001$] and RTs [$F1(1,17) = 11.16$, $p = 0.004$; $F2(1,62) = 9.67$, $p = 0.003$]. Hence a clear HF advantage, which is common to online word-processing tasks and thus highlights the validity of the experiment as well as the participants' sensitivity to L2 lexical cues.

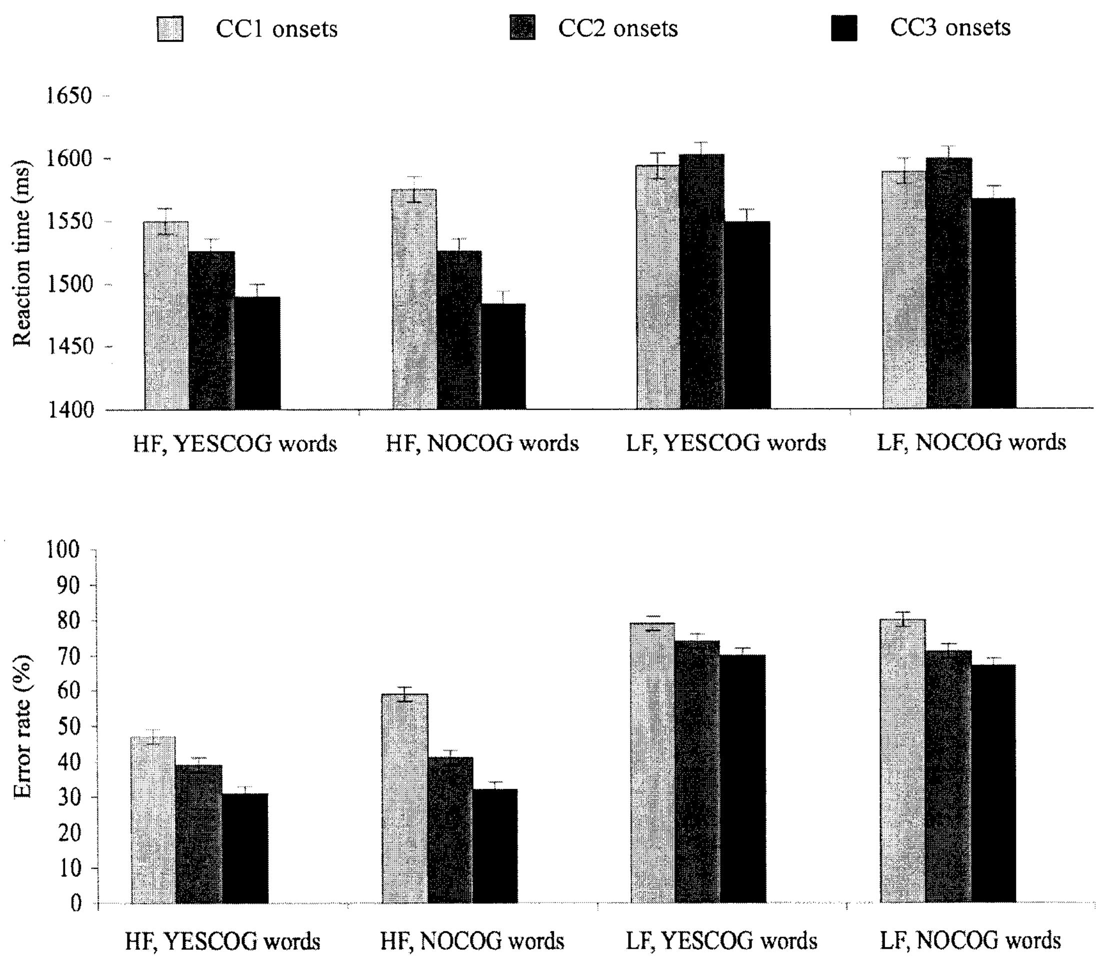


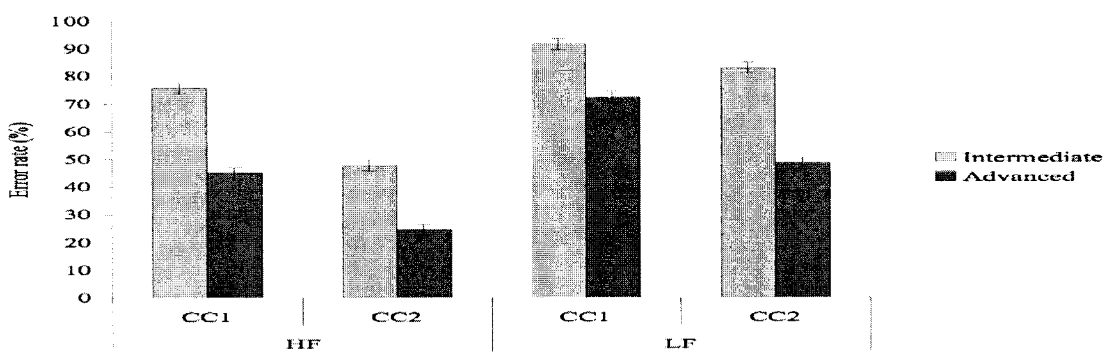

Figure 5. Word-spotting results of the Francophone participants (by stimulus variable)

*Onset Type*. Onset Type had significant effects on RTs [F1(2,34) = 5.36, p = 0.009; F2(2,22) = 3.87, p = 0.04] and error rates [F1(2,34) = 21.96, p = 0.0001; F2(2,22) = 9.15, p = 0.001]. Given the likelihood, confirmed in Figure 5, that Aligned strings (based on CC3 onsets) would be processed most accurately and fastest, the only planned by-subjects t-test here concerns the CC1 and CC2 conditions. For RTs, the CC1-CC2 contrast was not significant [t(35) = 1.82, p = 0.08], so that the only significant Onset Type difference opposed illegal (CC3) to legal (CC1, CC2) onsets. For error rates, however, the CC1-CC2 contrast reveals that CC1 onsets yielded significantly more errors than CC2 onsets [t(35) = 4.01, p = 0.0001]. In sum, therefore, word-spotting accuracy followed the ranking CC1 < CC2 < CC3, which matches the ranking of nonword processing accuracy in the lexical decision task: CC1 ($<$ C) $<$ CC2 $<$ CC3. Thus, whether dealing with nonwords or nonsense carrier strings, the Francophone ESL learners seemed to have used phonotactic cues in the same manner: the less French-like the onsets seemed (from the CC1 to the CC3 category), the more accurately they processed nonsense items based on these onsets.

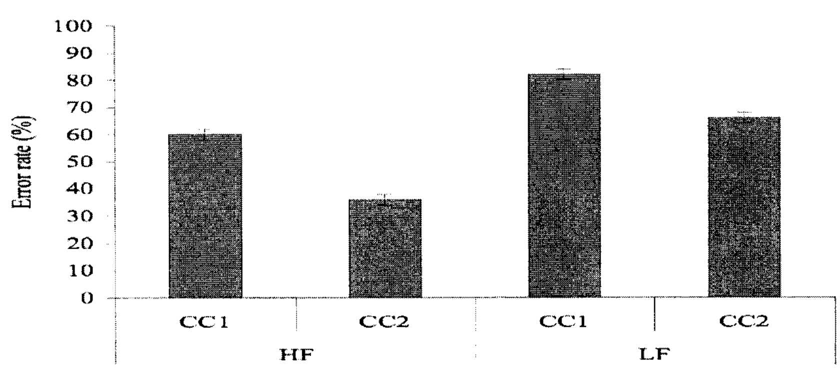
*Frequency x Onset Type*. The Frequency x Onset Type interaction was non-significant for RTs [$F1(2,289) = 1.49$, $p = 0.23$; $F2(2,269) = 0.45$, $p = 0.64$] but was weakly significant for error rates [$F1(2,289) = 3.46$, $p = 0.03$; $F2(2,269) = 2.37$, $p = 0.10$] in a by-subjects analysis. This result is surprising because planned Bonferroni-adjusted by-subjects t-tests (per test alpha of 0.025) showed that participants made significantly less errors with CC2 onsets than CC1 onsets not only in the context of HF words [$t(35) = -8.28$, $p = 0.0001$] but also in the context of LF words [$t(35) = -6.28$, $p = 0.0001$]. These parallel results (highlighted in the bottom graph of Figure 6) not only confirm that phonotactic "L1-unlikeness" and high L2 word frequency both act as segmentation cues, they also suggest that these phonotactic and lexical cues complement rather compensate (i.e., interact with) each other. In short, the above Frequency x Onset Type effect appears to be inconsistent with the lack of significant interaction from these t-tests.

A possible explanation comes from the higher-order, by-subjects interaction Frequency x Onset Type x Proficiency. This interaction was insignificant yet marginal [$F1(2,289) = 2.66$, $p = 0.07$]. This marginality could have been caused by a lack of power, itself due to a shortage of participants. In any case, it suggests a Proficiency-related trend, which may explain the interaction effect found above when Frequency and Onset Type were collapsed over Proficiency groups. This reasoning led to comparing the two graphs of Figure 6: the upper graph, showing mean error rates before the collapse over Proficiency groups; the lower graph, showing error rates after the collapse.

The upper graph reveals that in the HF condition, and regardless of Proficiency, learners made less word-spotting errors when dealing with phonotactic L1-L2 contrasts (CC2 condition) than when dealing with phonotactic L1-L2 similarities (CC1 condition). The LF condition remains a good measure of such CC2 benefits for Advanced learners, but it reflects CC2 benefits less clearly with regard to Intermediate learners. Indeed, there seems to have been an overall floor effect, whereby Intermediate learners might have benefited from phonotactic L1-L2 contrasts when processing HF words (i.e., words that they most likely knew) but not when processing LF words (i.e., words that they most likely ignored). This presumed floor effect could explain the interaction found earlier when error rates were collapsed over Proficiency groups. Specifically, in light of the upper graph, this floor effect could have inflated the Intermediates' error rate in the LF-CC2 condition, thus reducing the magnitude of the difference between the Intermediates' LF-CC1 and LF-CC2 error rates. Consequently, in the lower graph, the difference between LF-CC1 and LF-CC2 appears to be flattened out and highly comparable to the difference between HF-CC1 and HF-CC2: a graphical outcome that echoes the lack of interaction emerging between Frequency and Onset Type in the above t-tests.




Frequency x Onset Type effects broken down by Proficiency groups




Frequency x Onset Type effects collapsed over Proficiency groups

Figure 6. Frequency by Onset Type word-spotting effects on the Francophone participants

To conclude this digression, one may reiterate that phonotactic "L1-unlikeness" and high L2 word frequency both provide segmentation cues, which complement rather than compensate (i.e., interact with) each other. However, given the above analysis of Frequency x Onset Type x Proficiency effects, it seems more cautious not to specify this conclusion by contrasting the performances of Advanced vs. Intermediate learners.

Following the analytical scheme introduced in the last experiment, other potential Proficiency effects will nonetheless be tested and discussed in Section 4.3.4.3.

*Cognacy.* Cognacy, in Experiment 3, had somewhat impeded L2 lexical decision–with NOCOG words being responded to faster than YESCOG words. Here, Cognacy had no significant main effects [error rates: $F1(1,17) = 1.26$, $p = 0.28$; $F2(1,63) = 0.03$, $p = 0.86$; RTs: $F1(1,17) = 0.87$, $p = 0.37$; $F2(1,44) = 0.27$, $p = 0.60$].

Cognacy x Frequency effects on error rates were significant by subjects [$F(1,17) = 4.79$, $p = 0.04$; $F2(1,269) = 2.38$, $p = 0.12$]. Yet no clear influence of Cognacy is revealed by this result nor by the post-hoc Bonferroni-adjusted, by-subject t-tests (per test alpha of 0.01). These tests showed that: 1) for a set level of Cognacy, Frequency significantly and predictably affected accuracy scores [YesCog-HF vs. YesCog-LF: $t(35) = -11.52$, $p = 0.0001$; NoCog-HF vs. NoCog-LF: $t(35) = -8.02$, $p = 0.0001$]; 2) for a set level of Frequency, however, Cognacy had no significant effects [HF-YesCog vs. HF-NoCog: $t(35) = -1.95$, $p = 0.06$; LF-YesCog vs. LF-NoCog: $t(35) = 0.71$, $p = 0.48$].

This overall lack of Cognacy effects can be explained in at least three ways.

A stimulus-based explanation could be that the Francophone learners did not find the 'franglais' YESCOG words (e.g., 'must' or 'loft') particularly French-like in comparison with the more exclusively English NOCOG words (e.g., 'kite' or 'pull'). Recall that these 'franglais' loanwords had fulfilled the need for lexical materials that would not just display cognacy or non-cognacy with French, but also high or low usage frequency in English and a narrow inventory of onset consonants (i.e., [k, l, m, n, p, r]). Indeed, beyond the possibility that loanwords might have failed to seem L1-like, it could be that the YESCOG vs. NOCOG conditions had far lesser chances of eliciting contrastive lexical effects than the sharply distinct HF vs. LF conditions.

A task-based explanation is that semantic similarities between cognates, while possibly inhibitory when target words were played in isolation (in Experiment 3), might have been neutralized in the present experiment with nonsense stimuli.

An empirical explanation is that Cognacy, compared to Frequency and Onset Type, may be less helpful (if at all) for the recognition/segmentation of string-embedded L2 words. This explanation is consistent with eye-tracking evidence that Cognacy tends to inhibit the recognition/segmentation of string-embedded L2 words (Section 2.2.1.4).

*Test of extraneous compound-related lexical cues.* Each of the 96 target words was tagged for the following data, as computed by Celex for bisyllabic English compounds: 1) Compound Family Size, i.e. the number of compounds containing the target (e.g., 'coat-tail' or 'raincoat'); 2) Compound-Final Family Size, i.e. the number of compounds ending with the target (e.g., 'raincoat'); 3) Compound Family Frequency, i.e. the sum of the frequencies of the compounds containing the target; 4) Compound-Final Family Frequency, i.e. the sum of the frequencies of the compounds ending with the target.

Pearson product-moment correlation coefficients were used to test the relationship between these compound-related data and across-item mean error rates and RTs for the 96 target words. Fisher r-to-z transforms were then used to compare the results of these tests between the Intermediate and Advanced Proficiency groups.

Significant correlations emerged, notably concerning error rates (four top-left corner cells of Table 14). But these correlations were weak ($0.06 < r^2 < 0.10$) and scatter diagrams showed them as non-linear and inconsistent. Also, as shown in Table 15, no significant correlation contrast appeared between the two Proficiency groups.

In sum, uncertainties remain about the possibility of a correlation between the tagged Celex compound data and the observed word spotting responses. One may therefore reiterate the view (put forward in the discussion of stimulus control measures) that extraneous lexical variables other than Frequency or Cognacy would not have greatly influenced L2 speech segmentation in this word-spotting experiment. Moreover, the above evidence of significant Onset Type effects confirms that this experiment had tapped processes of speech segmentation rather than compound processing.

Table 14. Pearson correlations of word-spotting responses with compound family sizes

| | Error rates | | | Reaction Times | | |
|---|---|---|---|---|---|---|
| | Inter-mediate | Adv-anced | Difference (z-value) | Inter-mediate | Adv-anced | Difference (z-value) |
| Compound Family Size | -0.29* | -0.25* | -0.29 | -0.02 | -0.04 | 0.14 |
| Compound-Final Family Size | -0.32* | -0.30* | -0.15 | -0.09 | -0.02 | -0.75 |
| Compound Frequency | -0.15 | 0.03 | -1.23 | -0.05 | -0.05* | 0 |
| Compound-Final Frequency | -0.28* | -0.18 | -0.72 | 0.08 | -0.08 | 1.09 |

* $p < 0.05$

4.3.4.3. Results of the Francophone participants: Effects of Proficiency

Error rates are ignored as they helped build the Proficiency groups (Section 4.1.2.1).

*Main Proficiency effects.* Figure 7, by showing overall faster responses for Advanced learners than for Intermediate learners, confirms the possibility that L2 speech segmentation efficiency may grow with L2 proficiency. Yet it should also be noted, in the present case, that Advanced learners were only significantly faster than Intermediate learners on a by-items basis [$F1(1,17) = 1.92$, $p = 0.18$; $F2(1,269) = 57.74$, $p = 0.0001$].

*Proficiency x Frequency effects.* These interaction effects on RTs were not significant [$F1(1,289) = 0.50$, $p = 0.48$; $F1(1,269) = 0.05$, $p = 0.82$]. Nor were the Proficiency x Frequency x Onset Type effects [$F1(2,289) = 0.36$, $p = 0.70$; $F2(2,269) = 0.42$, $p = 0.66$]. Thus, Advanced learners did not gain significantly more from the effects of high word frequency, even when those were combined with the facilitatory effects of CC3 onsets.

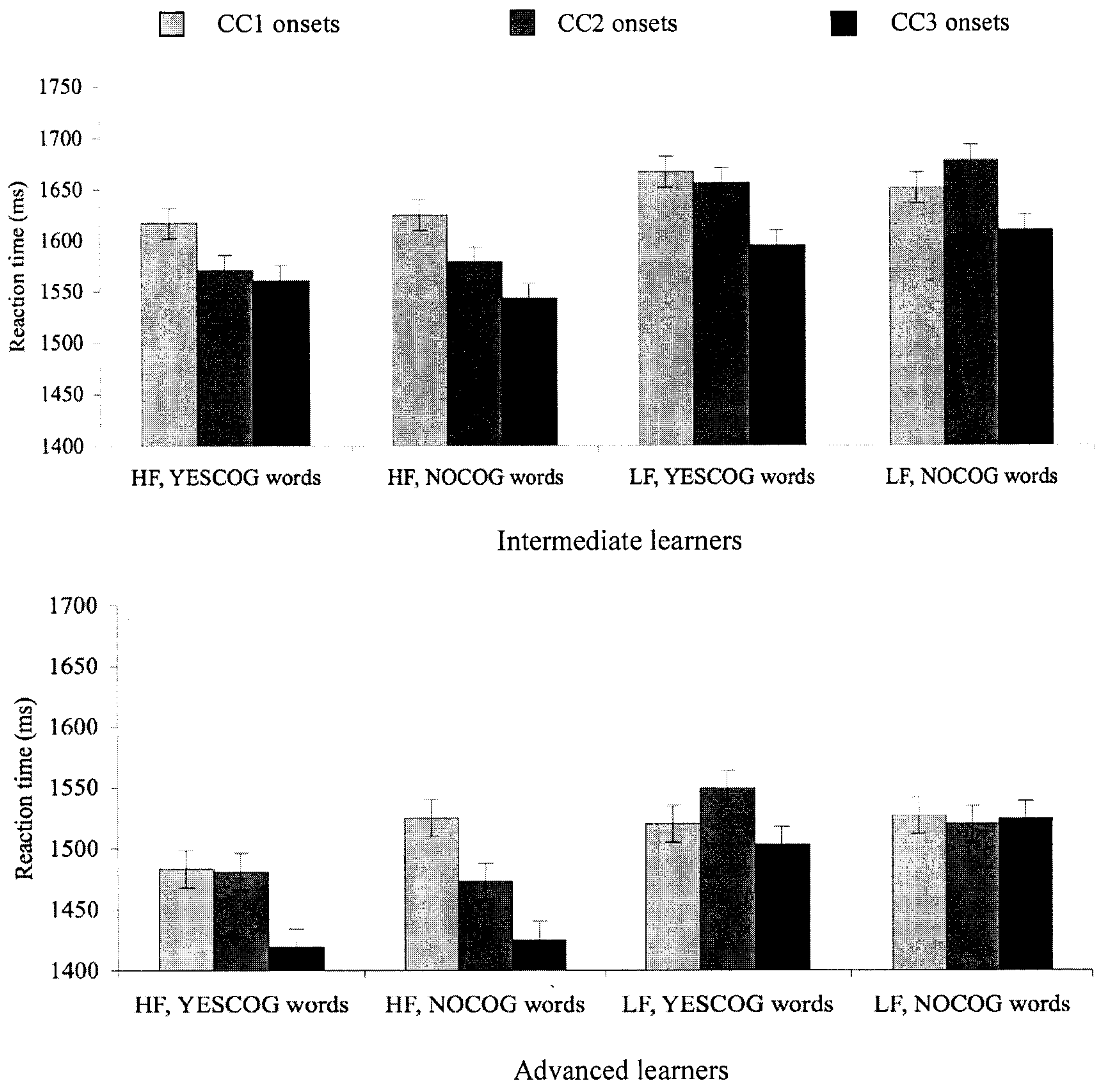



Figure 7. Word-spotting RTs of the Francophone participants (by Proficiency group)

*Proficiency x Onset Type effects.* The Proficiency x Onset Type interaction for RTs was not significant [$F1(2,289) = 1.29$, $p = 0.28$; $F2(2,269) = 0.35$, $p = 0.71$].

Further analysis considered how the Proficiency groups compared with respect to sonority factors in the processing of Aligned (CC3-based) strings. In the pilot study, previous analyses of this kind had classified CC3 onsets as 'tolerable', 'bad' and 'very bad', depending on their statuses vis-à-vis the Sonority Sequencing (SD) and Minimal

Sonority Distance (MSD) constraints. Here, only two onset subtypes applied: 'tolerable' (fk, fp]) and 'very bad' ([km, pn, sr, tl]).

Sonority effects appear more clearly than during the lexical decision task, with 'very bad' onsets being rejected faster than 'tolerable' ones (Figure 8). This is especially true for Advanced learners, though the Proficiency x CC3 Subtype effects on RTs were only significant by items [F1(1,17) = 2.70, p = 0.12; F2(1.15) = 4.86, p = 0.04].

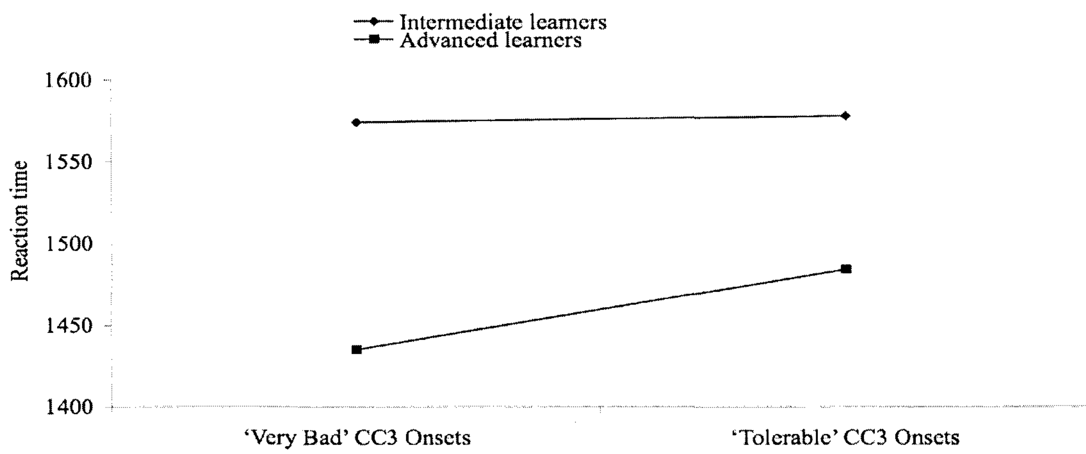



Figure 8. Effects of sonority on learner performance in the word-spotting task

These last results suggest that phonotactic-based L2 speech segmentation involves sonority considerations, perhaps all the more so as learners gain L2 proficiency.

4.3.4.4. Summary and conclusions

*A cue strength hierarchy.* Such a hierarchy seems to emerge from Experiments 3 and 4.

Frequency can be thought to dominate this hierarchy, given that it significantly affected the processing of isolated targets as well string-embedded targets.

Onset Type/Phonotactics can be thought to occupy a second position in this hierarchy, given its mitigated effects on lexical decision (i.e., better recognition of words with native-like C/CC1 onsets than with non-native-like CC2 onsets) and its mitigated effects on word spotting (i.e., better word recognition after non-native-like than after native-like string-medial onsets).

Cognacy can be thought to lie at the bottom of this cue strength hierarchy, though its status remains to be specified. The conclusion may be that it is a non-cue if the only Cognacy-related pattern emerging from the study is a hampering effect on lexical decision RTs. Or one may conclude that it is a weak cue if the next, offline segmentation task (Experiment 5) shows that Cognacy can in fact affect L2 speech segmentation.

*Phonotactics - recapitulation and discussion.* Four findings are relevant here.

First, L2 listeners were phonotactically sensitive, as shown by their significantly better results in the Aligned (CC3) condition than Misaligned (CC1/CC2) condition.

Second, the lack of a clear interaction between Frequency and Onset Type hints that phonotactics may act independently of, rather than as back-up for word frequency.

Third, 'very bad' CC3 onsets were stronger segmentation cues than 'tolerable' CC3 onsets, suggesting that L2 phonotactic segmentation might involve sonority factors.

Fourth, CC1 onsets appeared to be most detrimental to L2 speech segmentation. That CC1 onsets should cause word-spotting difficulties is hardly surprising. Word-spotting studies of L1 French (Dumay, Content & Frauenfelder, 1999, 2000, 2002) had already shown the inhibitory effects Misaligned strings built on such onsets (e.g., [bl] or [tʁ]), contrasting with the facilitatory effects of Aligned strings built on illegal (CC3) onsets[28]. The question is why, in L2 English segmentation by Francophone listeners, these L1-like stop-liquid onsets (or "CC1 onsets", in this study) turn out to be more detrimental than L2-like [s]-stop onsets like [sp] or [st] ("CC2 onsets").

A general explanation, legitimized by the non-effects of Cognacy, is that L2 segmentation draws efficiency from the non-transfer of L1 phonotactic and lexical cues.

An alternative, but potentially complementary, explanation involves the possible influence of L1 sublexical patterns on L2 speech segmentation. This explanation emerges from two previous studies, as reviewed below.

[28]For Dumay et al. (2002: 148), stop-liquid clusters are best suited to build Misaligned strings because they are the only clusters unanimously considered in the French phonology literature as non-separable.

Beaudoin (1996) found that, in line with the theoretical predictions of Delattre (1940) and Dell (1995), French-dominant, French-English bilinguals tended to syllabify intervocalic, stop-liquid/CC1 clusters as onsets, but tended to split [s]-stop/CC2 across a syllable boundary rather than preserve them as onsets. These differential patterns provide behavioral support for the statistical distinction made in this study between stop-liquid/CC1 onsets (unmarked in French) and [s]-stop/CC2 onsets (marked in French). These patterns, moreover, can be deemed capable of affecting word-spotting in English, so that Francophone ESL learners would treat CC1 onsets as more acceptable, hence more characteristic of the Misaligned scenario (as shown by a segmentation response like vu.f#last), and ultimately more detrimental than CC2 onsets (as in vus.#case).

Dumay, Content and Frauenfelder (1999) further evidenced the different statuses of stop-liquid/CC1 and [s]-stop/CC2 onsets in French. Their second word-spotting task was designed by splice-recording CC1 or CC2 onsets and accompanying carrier strings from two types of French source phrases: source phrases containing the same noun as the target noun in the carrier vs. source phrases containing a resemblant noun (see Table 15).

Table 15. Onset type variations in Dumay, Content & Frauenfelder (1999)

| Status of target noun (betw. carrier string & source phrase) | Onset type | |
|---|---|---|
| | CC1 onset<br>Ex.: [tʁ]<br>String: [tɑ̃.t#ʁu] - Target: *tante* | CC2 onset<br>Ex.: [st]<br>String: [ʁa.s#ty] - Target: *race* |
| Same | [tɑ̃t#ʁublard]<br>'*tante* roublarde' ('crafty aunt') | [ʁas#tymefie]<br>'*race* tuméfiée' ('swollen race') |
| Resemblant | [tɑ̃#tʁublɑ̃]<br>'*temps* troublant' ('troubling time') | [ʁa#stypefɛ]<br>'*rat* stupéfait' ('stupefied rat') |

For CC1 onsets, word-spotting appeared to be faster in the Same than in the Resemblant condition[29], and subsequent acoustic analyses of the source phrases revealed systematic lengthening of the liquid and vowel preceding the word boundary. CC2 onsets, in contrast,

[29]The authors had in fact named these conditions Aligned vs. Misaligned. For clarity, they are renamed here Same vs. Resemblant. Indeed, in the present study, Aligned and Misaligned refer to a relationship between word and syllable boundaries within a nonce string rather than between a nonce string and a source phrase.

showed no word-spotting facilitation or segmental lengthening. The authors concluded that phonetic/durational cues were easier to detect on stop-liquid (CC1) onsets than on [s]-stop (CC2) onsets, and that L1 French segmentation drew upon such cues. Back to the present study, and given the nascent state of research on phonetic cues to L2 speech segmentation, one cannot exclude the possibility that L1 French listeners might transfer the expectation of such durational, French-specific patterns to L2 segmentation in English. It would follow, in compliance with the CC1 < CC2 scale of accuracy found in Experiments 3 and 4, that: 1) CC2 onsets, seeming phonotactically legal in English, should be detrimental to L2 English segmentation; 2) CC1 onsets, seeming both phonotactically legal and phonetically variable, should be even more detrimental to L2 English segmentation. Such a conclusion, of course, could only be fully validated by applying Dumay et al's mix of word-spotting tasks and phonetic/durational analyses (in both L1 and L2) to the case of Francophone ESL learners.

In sum, the lower costs of CC2 over CC1 onsets, as found in the present experiment, could be explained by native syllabification patterns (i.e., the fact that CC2 clusters are less onset-like) and/or native durational patterns. Either way, the L2 learners still ended up being misled by CC2 onsets, as they were expected to be under the overall Misaligned condition. In other words, potentially useful, sublexical L1 information seems to have given limited benefits.

On this basis, an informal, two-step resolution of the "L1 transfer" issue can be attempted. It draws from the above internal and external findings, and exploits the fact that cognacy, preferrable onset layouts and systematic durational variations respectively constitute lexical, syllabic and phonetic patterns. This informal conclusion reads thus:

- transfer (of lexical, syllabic and phonetic patterns) achieves little.
- non-transfer (of lexical and syllabic patterns) seem more worthwhile.

*Left-over issues (for Experiments 5 and 6).* Experiment 5 examines whether word-spotting performance might have been inadvertently affected by semantic noise resulting from the presence of real words inside the carrier strings. Experiment 6 examines the strength of the phonotactic biases found in the above word-spotting task.

Experiments 5 and 6 are reported in the next sections.

## 4.4. Experiment 5: Word-likeness judgment

### 4.4.1. Rationale

In Experiment 4, the lack of significant correlations between compound-related data and word-spotting responses strengthened the view that extraneous lexical variables other than Frequency or Cognacy had not been influential. There still remained a possibility that word-spotting performance, originally supposed to tap the retrieval of phonological lexical entries, could have been inadvertently affected by lexico-semantic noise resulting from the presence of real words inside the stimuli.

The aim of the present experiment was to assess the possibility of such noise in Experiment 4. To this aim, participants were asked to rate the word-likeness (i.e., lexico-semantic plausibility) of the carrier strings used in the word-spotting task.

### 4.4.2. Materials

The 96 carrier strings from the word-spotting task (Experiment 4) were reused.

### 4.4.3. Procedure

The participants were tested one by one. An iMac running PsyScope 1.2.2 was used to present stimuli and collect word-likeness ratings. Each trial consisted in a 500 ms warning asterisk followed by a headphone presentation of the stimulus.

English instructions on the iMac screen instructed participants to rate the items from Experiment 4 as '1. impossible', '2. unlikely', '3. likely', or '4. very likely', depending on how they wished to answer the question "could this item pass for a new word of English?". Participants had to press 1, 2, 3 or 4 on the keyboard to indicate their favored rating. Four practice trials were given and later excluded from the analysis.

### 4.4.4. Results and discussion

4.4.4.1. Preliminary statistics

The above rating scheme can be rephrased as four steps on a word-likeness scale: '1. not word-like at all', '2. not word-like', '3. word-like', '4. very word-like'.

The mean rating was 2.4. There is little risk that this result may have derived from a deliberate response strategy (e.g., 'keep pressing 1 or 2') since all four ratings were fairly uniformly covered (as shown by the respective proportions of 28%, 27%, 27% and 19% for the 1, 2, 3 and 4 ratings).

Figure 9 above confirms this widespread distribution.

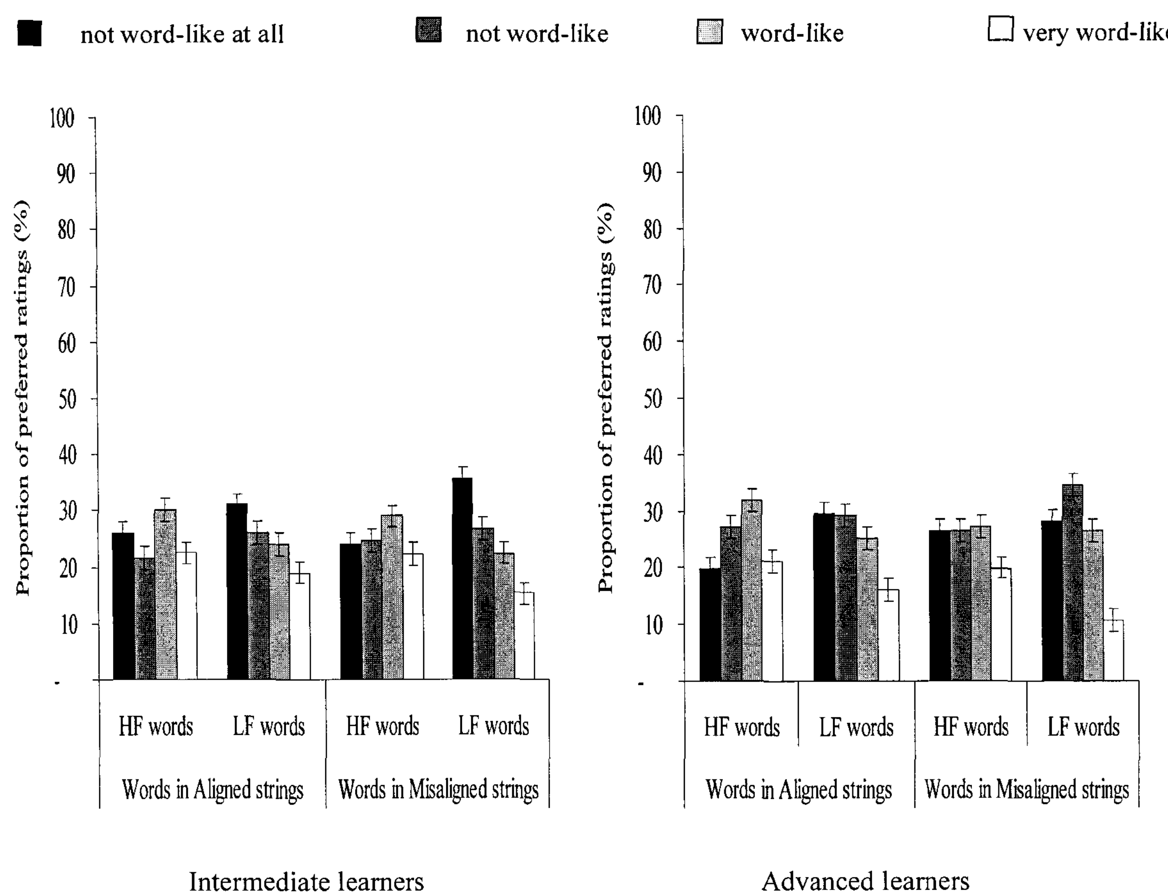



Figure 9. Word-likeness ratings of stimulus strings by the Francophone participants

Figure 9 illustrates two main response patterns shared by Intermediate and Advanced learners as well as by the Aligned condition (CC3 Onset Type) and the Misaligned condition (CC1 and CC2 Onset Types). First, carriers ending in HF words (e.g., 'vooc.man') received more 'word-like' ratings. Second, carriers ending in LF words (e.g., 'thouk.mule') received more 'not word-like' or 'word-like at all' ratings. The significance of these patterns is explored next.

#### 4.4.4.2. Significance testing

In preamble, it should be pointed out that the collected data (from '1. not word-like at all' to '4. very word-like') are ordinal. As such, they should not lend themselves to 'interval-type statistics' like ANOVAs or t-tests. Yet the 'rating' variable used in this task can be said to contain some interval information in that it measures the magnitude of a specific attribute (i.e., the perceived word-likeness of a nonensical bisyllable). Indeed, the past decades of statistical research have seen a growing leniency for the parametric testing of ordinal data (cf. Chen, Gusshoven & Rietveld, 2004 for review). It is on these grounds that the word-likeness ratings were averaged and that ANOVAs and t-tests were run for this experiment as they had been for the previous experiments.

In an initial ANOVA involving averaged word-likeness ratings and incorporating the variable Test List (along with Onset Type, Cognacy, Frequency and Proficiency), no effects of Test List was found (all Fs $< 1.89$, p. $> 0.17$). Rating data were therefore collapsed across test lists in the new ANOVA Onset Type x Cognacy x Frequency x Proficiency. Here are the main results of this ANOVA[30].

*Onset Type*. Onset Type showed no significant effects on word-likeness ratings [$F1(2,34) = 2.20$, $p = 0.13$; $F2(2,22) = 1.15$, $p = 0.33$].

*Cognacy*. Cognacy showed no significant effects on word-likeness ratings [$F1(1,17) = 0.13$, $p = 0.73$; $F2(1,35) = 0.19$, $p = 0.67$].

[30]All variable interactions were insignificant (all Fs $< 2.47$, p. $> 0.09$).

*Proficiency.* Proficiency showed no significant effects on word-likeness ratings [$F1(1,17) = 0.04$, $p = 0.84$; $F2(1,93) = 0.61$, $p = 0.44$].

*Frequency.* Frequency showed significant effets [$F1(1,17) = 22.22$, $p = 0.0001$; $F2(1,44) = 8.87$, $p = 0.005$]. This being said, the 2.5 and 2.2. ratings that were respectively found for the HF and LF categories remain within the range of 'not word-like' ratings.

4.4.4.3. Conclusions

Participants, regardless of their level of L2 proficiency, did not overall feel that the carrier strings from the previous, word-spotting experiment were word-like–as reflected by an overall mean rating of 2.4 for the Intermediate learners and of 2.3 for the Advanced learners. This finding supports the view that extraneous lexical variables may not have corrupted the results of the previous, word-spotting experiment.

This conclusion is unaffected by the above frequency effects, since both HF and LF items received mean ratings corresponding to the label 'not word-like' (respectively 2.5 and 2.3 for carriers of HF and LF words). The fact that carriers of HF words were perceived as slightly more word-like simply endorses the earlier choices of HF vs. LF stimuli, that is, of materials capable of triggering the sensitivity of ESL learners to frequency variation in both online and offline experiments.

In sum, this control, offline experiment has given a further validation of the word-spotting task.

Before recapitulating the outcome of the overall study, it nonetheless seemed useful to run a final check of how responsive the Francophone ESL learners were to English phonotactics. This final check was the aim of Experiment 6, as reported next.

## 4.5. EXPERIMENT 6: OFF-LINE SEGMENTATION OF PRINTED CVCCCVCs IN L2 ENGLISH

### 4.5.1. Rationale

Experiment 4 had shown, on the basis of nonsense CVCCVC(C) strings ending in CVC(C) words (such as 'voofpose' or 'voospose'), that word-spotting by Francophone ESL learners was affected by the phonotactics of the string-medial CC onsets, according to the accuracy ranking CC1 < CC2 < CC3. Thus, the less French-like these clusters (from onset type CC1 to onset type CC3), the more efficiently participants split these clusters across a syllable boundary and spotted the string-final CVC(C) words as a result.

Experiment 6 retested this pattern offline to check if participants were indeed sensitive to English phonotactics and to the transferrability of French phonotactics. The visual mode was adopted, as in Experiment 1, to bypass the risk of cluster mishearings.

### 4.5.2. Materials

Eighty CVCCCVC strings (e.g., 'loufrich' or 'voobleck') were used. Half were carrier strings reused from Experiment 4 (e.g., 'loufrich'); half were newly designed, with no embedded word (e.g., 'voobleck'). The overall set (Appendix K) allowed the association of 4 strings (i.e., 2 carrier strings from Experiment 4 plus 2 newly designed strings) to each of the clusters tested in Experiment 4 (i.e., the CC1 onsets [bl, fl, gl, kl, pl, br, fr, gr, kr, pr], the CC2 onsets [sk, sm, sn, sp], and the CC3 onsets [fk, fp, km, pn, sr, tl]).

### 4.5.3. Procedure

Materials were shown on a printed sheet. Participants had to indicate what they viewed as the best segmentation for each CVCCCVC string by drawing a slash in the desired position. Thus, if faced with 'voobleck', they could segment graphically as 'voob/leck' or 'voo/bleck'. Two random orders of presentation were assigned.

### 4.5.4. Results and discussion

4.5.4.1. Data coding

Transcriptions were discarded if the CC onset of interest was split across the slash sign (e.g., 'voob/leck'). They were kept for further statistical analysis if they showed the CC onset being preserved to the right of the slash sign (e.g., 'voo/bleck').

4.5.4.2. Statistical outcome

*Preliminary statistics.* The onset preservation ranking CC1 > CC2 > CC3 (Figure 10) is explored next, on the basis of an ANOVA of Onset Type, Proficiency and String Type (i.e., carriers from Experiment 4 vs. strings devoid of real words, or non-carriers) .

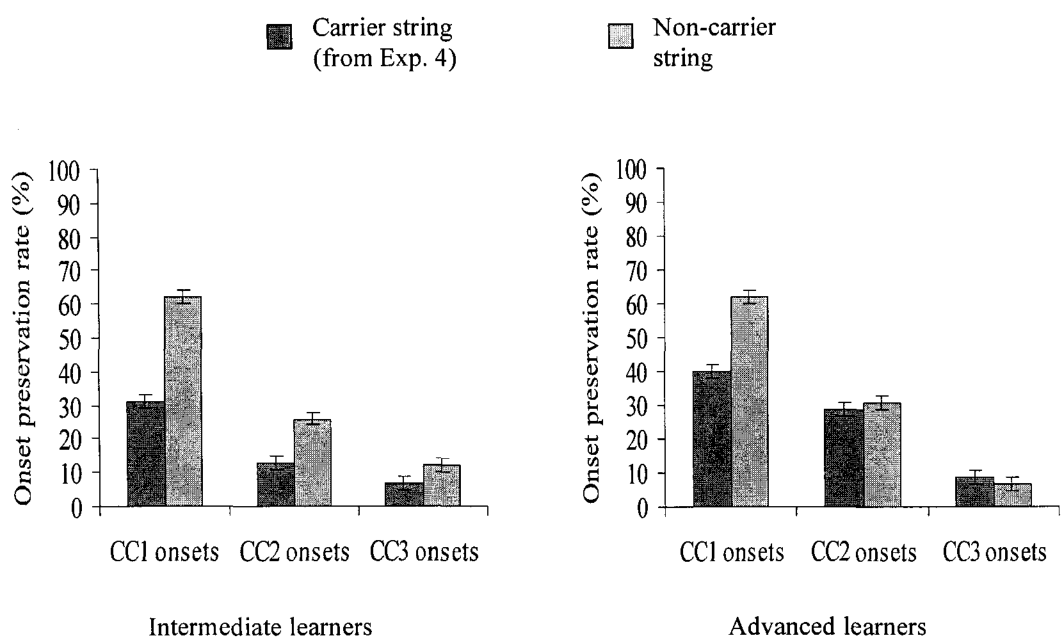



Figure 10. Preservation of printed onset clusters by the Francophone participants

*Onset Type.* Onset Type had significant by-subjects and by-items effects [$F1(2,34) = 101.70$, $p = 0.0001$; $F2(2,18) = 42.44$, $p = 0.0001$]. Bonferroni-adjusted by-subjects t-tests (per test alpha of 0.017) showed that onset preservation rates had been significantly higher for CC1 than CC2 [t(35) = 7.05; p = 0.0001], for CC2 than CC3 [t(35) = 7.17, p = 0.0001], and for CC1 than CC3 [t(35) = 13.31, p = 0.0001]. These results reflect the onset preservation ranking CC1 > CC2 > CC3 that emerges from Figure 10.

*String Type*. String Type had significant main effects, with higher preservation rates when the CC onsets of interest appeared in non-carrier strings than in carrier strings [F1(1,17) = 34.56, p = 0.0001; F2(1,23) = 16.12, p = 0.001].

String Type interacted significantly with Onset Type [F1(2,34) = 20.60, = 0.001; F2(2,92) = 25.55, p = 0.0001]. Bonferroni-adjusted t-tests (per test alpha of 0.017) showed that, while carriers and non-carriers did not yield significantly different CC3 preservation rates [t(35) = 1.65, p = 0.11], non-carriers yielded significantly higher rates than carriers with respect to CC1 onsets [t(35) = 7.63, p = 0.0001] and CC2 onsets [t(35) = 4.65, p = 0.0001]. This result is unsurprising since CC1 or CC2 onsets like [pl] or [sk] would be more likely preserved in items like 'chauploff', which do not end in real CVC(C) words and thus encourage segmentations like 'chau/ploff', than in items like 'vawbreal', which do end in real words and thus may be visually segmented via a lexical route ('preserve the word') as well as via a phonotactic route ('preserve the cluster').

*Proficiency*. The mean preservation rate was 30% for the Advanced learners, 25% for the Intermediate learners. This difference was only but highly significant in a by-items analysis [F1(1,17) = 0.88, p = 0.36; F2(1,92) = 10.37, p = 0.002], confirming the higher sensitivity of these Advanced L2 learners to L2 phonotactics. Proficiency did not significantly interact with Onset Type [F1(2,34) = 1.88, p = 0.17; F2(2,92) = 2.63, p = 0.08] or String Type [F1(1,17) = 0.01, p = 0.92; F2(1,92) = 0.05, p = 0.82].

4.5.4.3. Conclusions

The onset preservation ranking CC1 > CC2 > CC3 shows that, as in the word-spotting task, the participants were less inclined to preserve string-medial CC onsets as these CC onsets got less French like (from onset type CC1 to CC3).

# CHAPTER FIVE

# CONCLUSIONS

## 5.1. Recapitulation of the study

### 5.1.1. Motivations

The study stemmed from an observation common among learners, teachers and researchers: segmenting incoming L2 speech into words is a tricky, yet learnable skill.

The study also stemmed from the perception of a current research gap. Weber and colleagues having recently begun to probe L1 and L2 phonotactic effects on L2 speech segmentation, it remained to be seen how such effects might compare with L1 and L2 lexical effects. This issue seemed equally relevant from an online perspective (i.e., as speech unfolds) and from a developmental perspective (i.e., as L2 acquisition unfolds).

The study, in this context, sought to assess the contributions of L1 and L2 lexical and phonotactic cues in the performance and the acquisition of L2 speech segmentation.

### 5.1.2. Methods

The contributions of L1 and L2 lexical and phonotactic cues were assessed by means of online experiments (testing how L2 listeners processed isolated or embedded words in real time) and back-up, offline experiments (checking that the online stimuli were fit and that phonotactics had the potential of helping English segmentation).

The online experiments, namely lexical decision and word-spotting, can be seen as highly complementary, for three reasons: 1) they involved comparable online word processing skills; 2) they used the same participants, target words and target clusters; 3) they both replicated standard online patterns (whereby speed and accuracy increase with high-frequency and phonotactically aligned words). It is therefore against the findings of these two experiments that the hypotheses of the study will now be evaluated.

5.1.3. Evaluation of hypotheses

Hypothesis 1 (i.e., L2 lexical and phonotactic effects on L2 speech segmentation) is supported. For both Proficiency groups, word-spotting speed and accuracy ranked as: LF-Misaligned < LF-Aligned < HF-Misaligned < HF-Aligned (Aligned and Misaligned conditions referring respectively to CC1/CC2 onsets and CC3 Onsets). This ranking means that: 1) frequency helped, through shifts from LF to HF words, 2) phonotactics helped, through shifts from Misaligned to Aligned strings, 3) frequency helped more.

Hypothesis 2 (i.e., L1 lexical and phonotactic effects on L2 segmentation) is partially supported. Lexical-cognacy effects helped in neither of the main experiments. And phonotactics shared by the L1 and the L2 (in the form of CC1 onsets) has proven more detrimental to L2 English segmentation than more L2-specific phonotactics (in the form of CC2 onsets).

Hypothesis 3 (i.e., L2 proficiency effects on L2 speech segmentation) is also partially supported. On one hand, Advanced learners were faster than Intermediate learners in lexical decision. They were also faster and more sonority-driven in word-spotting, at least in the relevant by-items analyses. On the other hand, they were not more sensitive than Intermediate learners to Frequency or Onset Type variations. On this basis it is difficult to decide between the alternative developmental paths proposed under Hypothesis 3 (i.e., Subhypothesis 3a: gradual shift from phonotactic to lexical cues; Subhypothesis 3b: increasing integration of lexical and phonotactic cues).

To sum up, it is in connection with Hypotheses 1 and 2 that the more solid conclusions of this study can be drawn, namely:

1) L2 phonotactics helps; L2 word frequency helps more;

2) L1 cognacy and L1 phonotactics do not help;

## 5.2. DISCUSSION OF CONCLUSIONS

### 5.2.1. "L2 phonotactics helps; L2 word frequency helps more"

This pattern parallels the fact that L1 segmentation also involves cue integration, with a priority most likely given to the search for meaningful, lexical units in the input. In more formal terms, L2 speech segmentation seems accountable for by a speech processing model like Shortlist (Norris, 1994), which views the competition of lexical candidates as the impetus for the lexical segmentation of incoming speech, while also emphasizing that lexical competition may be boosted by the listener's access to sublexical cues.

This parallel between L1 and L2 segmentation procedures may itself be taken as a more general indication that human speech processing is first and foremost driven by the search for lexical meaning, with an additional and somewhat opportunistic interest in processing cues located below the word level[31].

### 5.2.2. "L1 cognacy and L1 phonotactics do not help"

*Rehabilitating L1-L2 contrasts.* Given the conclusion that L1 cognacy and phonotactics do not help, and given the L2 effects discussed above, it seems that L2 segmentation undergoes the following scale of increasingly helpful cross-linguistic influences:

| | | | | |
|---|---|---|---|---|
| *Properties (seemingly) shared by the L1 & L2* | | *Properties (seemingly) specific to the L2 only* | | *Properties (seemingly) universal* |
| Examples: | | Examples: | | Example: |
| 1. L1-L2 cognacy | < | 3. L2 word frequency | < *or* = | 5. Illegality of CC3 onsets |
| 2. Unmarkedness of onsets in both L1 & L2 (CC1 condition) | | 4. Unmarkedness of onsets in L2, not L1 (CC2 condition) | | 6. Felicity of 'very bad' over 'tolerable' onsets (CC3 condition) |

[31]This type of optimization procedure might be seen as reminiscent of Zipf's (1949) Force of Diversification in human listening–even though this concept was more specifically concerned with the idea that human listeners strive to attach one distinct meaning to each single word that they hear.

This scale is in line with theories challenging the orthodox view that L2 processing/acquisition is bound to gain from L1-L2 similarities and suffer from L1-L2 dissimilarities (e.g., Lado, 1957). In particular, the Moderate Contrastive Analysis Hypothesis (Oller & Ziahosseiny, 1970) and the Similarity Differential Rate Hypothesis (Major & Eunyi, 1999) respectively hold that dissimilarities are in fact easier and faster to learn than similarities. The Speech Learning Model also holds that the greater the perceived phonetic distance between L1 and L2 sounds, the easier for a learner to establish an independent phonetic category for the relevant target L2 sound (Flege, 1995). This proposal is especially relevant here as L2 segments and their phonotactic arrangements are likely to be learnt in tandem, through gradual exposure to L2 words made of these segments.

More generally, if properties universal or specific to the target language end up being more helpful than L1-L2 similarities, then one can reiterate the view that L2 listeners are perhaps not that different from L1 listeners in how they segment input.

*Psychotypological considerations.* 'Seemingly', in the above scale, refers to the fact that the assessment of speech segmentation cues (phonotactic cues, in particular) is inherently subjective. For any listener, whether phonotactic cases 2, 4, 5, 6 above refer to language-specific or universal patterns of markedness is indeed matter for subjectivity, given that the multiple phonotactic scenarios found across languages are not accessible to non-native intuitions. Consequently, learners being asked to assess the phonotactics of a given CC cluster seem bound to rely on more or less accurate impressions.

Such subjectivity may play a role as L2 phonotactics gets processed online. This possibility is sustained by the psychotypological theory of SLA, whereby L1 transfer may operate on the basis of the learners' conscious as well as unconscious perceptions of L1-L2 distance (Kellerman, 1983). Moreover, Experiment 6 has revealed patterns of legal onset preservation that were not just significant but also consistent with the online results of Experiment 4, as if the learners' subjective, offline assessments of L2 phonotactics could affect their online use of phonotactics in L2 segmentation. This scenario being thus both theoretically and empirically viable, it has an immediate pedagogical implication (further discussed in Section 5.4): by explicitly teaching the phonotactics of an L2, one may arm learners with a knowledge that could efficiently (though unconsciously) help them to segment L2 speech input in real time.

*The non-effects of Cognacy.* This finding has already been explained in both methodological terms (i.e., failure of stimuli to elicit a sharp contrast between cognacy and non-cognacy with French; neutralization of semantic cognacy effects in word-spotting) and empirical terms (i.e., eye-tracking evidence that Cognacy tends to inhibit L2 segmentation). A theoretical explanation can be attempted at this conclusive stage.

The non-effects of Cognacy should first be differentiated from a non-cognate advantage, which would have meant consistently faster performance with NOCOG words than YESCOG words. This pattern did occur in the lexical decision task, but not in the word-spotting task. In the absence of such a consistent non-cognate advantage, the findings of this study do not clearly favor the 'non-selective language view' of bilingual lexical activation (i.e., concurrent processing of L1 and L2 representations).

The non-effects of Cognacy could in fact fuel the opposite scenario, whereby the participants, prior to the study, may have gained enough proficiency in L2 English for their L1 and L2 lexical representations/processes to be separate. The impression of an overall high L2 proficiency in these participants comes from qualitative evidence that they were far from beginners in ESL training, and from quantitative evidence that, in both word-spotting and lexical decision (for words and nonwords), Advanced learners had been significantly faster than Intermediate learners on a by-items basis but not on a by-subjects basis. As for the possibility of a gradual split of languages in a bilingual's lexicon, it is conceivable if one assumes that over the course of L2 development, the links between L2 concepts and L2 word forms might grow ever stronger. This evolution could lead to a stage where L2 concepts and L2 word forms get separate from their L1 counterparts so that, ultimately, spreading of lexical activation across the two languages (evidenced by facilitatory cognacy effects) may become difficult (see Woutersen, Cox, Weltens & de Bot, 1994 for a similar argument).

On a level more closely related to the functioning of L2 speech segmentation per se, one may here reemphasize that this lack of facilitatation by L1-like words parallels the inhibitory effects of L1-like, CC1 onsets. This parallel clearly suggests, as pointed out in the earlier discussion of the word-spotting results, that L2 segmentation may generally gain from the non-transfer of L1 cues, be they lexical or phonotactic.

*The costs of shared L1-L2 phonotactics.* This discussion compares the findings and interpretations of the present study with those of Weber and Cutler (2006) using German ESL learners. For this comparison, the terms "cue" and "non-cue" respectively refer to onset clusters pertaining to the Aligned and Misaligned word-spotting conditions.

A taxonomy of cues and non-cues emerges from these combined studies:

a. "L1-L2 cues" (i.e., onsets that are illegal in both L1 and L2, like [kp] or [ml]);
b. "L1-L2 non-cues (i.e., legal onsets of comparable L1 and L2 frequencies, like [pl]);
c. "more L2-like non-cues" (i.e., legal onsets of higher L2 frequency, like [sp]);
d. "L1 non-cues (i.e., legal onsets in the L1 only, like [ʃl]);
e. "L2 non-cues" (i.e., legal onsets in the L2 only, like [sl]).

The studies find that, against the baseline category a., and in comparison with categories b. and c. (as in this study) or d. and e (as in Weber and Cutler's)[32], "L1-L2 non-cues" are particularly detrimental to L2 speech segmentation. Hence, from both studies, an indication that phonotactics shared by the L1 and the L2 can be costly.

This common finding strengthens Weber and Cutler's view that "proficient L2 listeners can acquire the phonotactic probabilities of an L2 and use them to good effect in segmenting continuous speech, but at the same time they may not be able to prevent interference from L1 constraints in their L2 listening" (p. 597). Yet, given the earlier discussions of this chapter, this statement would better fit the present research if the phrases "proficient L2 listeners" and "L1 constraints" were rephrased to the more specific "non-beginner L2 listeners" and "L1 phonotactic constraints".

---

[32]In Weber and Cutler's paper, a result graph for the German learners (as opposed to the control native Anglophones) shows that "L1 non-cues" (or Condition 2, in Section 2.2.2.4 of the present literature review) generated faster but less accurate responses than "L2 non-cues" (or Condition 3 in that same review). The authors however do not report a statistical test of this contrast.

### 5.2.3. Final interpretations

Unlike Weber and Cutler (as quoted above), and despite the lack of evidence for a stronger use of lexical and phonotactic cues by Advanced learners, one may conclude this study by emphasizing the learnability of L2 segmentation skills over the risk of L1 influences on these skills.

This preferred emphasis on learnability over interference draws on the following key findings: native-like prioritization of lexical over phonotactic cues; non-effects of interlingual cognacy; indications (albeit weak, since only reliable by items) that learners might get faster and more sonority-driven[33] as they advance in the L2.

Another important finding is that Intermediate L2 learners were also using phonotactic cues associated with English. This finding suggests that, at early stages in L2 acquisition, learners are ready to exploit linguistic information specifically useful for the purpose of L2 speech segmentation. This suggestion, along with the possibility for L2 segmentation to become L1-like, reinforces the view expressed at the outset of this dissertation, that L2 speech segmentation may be of key importance to L2 acquisition.

## 5.3. Directions for further research

### 5.3.1. Possible methodological improvements

*Pool changes*. The study did not find clear evidence that cue integration might increase along with L2 proficiency. This may be because advanced learners were compared with intermediate learners. An equivalent study, therefore, would be worth running with absolute beginners. Actually, in light of the consensus whereby positive L1 transfer typically occurs at low-proficiency levels (e.g., Odlin, 1989; Poulisse & Bongaerts, 1994), the conclusion that L1 lexical/cognacy cues helped little may have to do with the

[33]Should this pattern be replicated and strengthened elsewhere, it could be explained in theoretical terms (i.e., the possibility for L2 lexical acquisition to modulate L2 phonological representations - cf. James, 1988: 88-91) and empirical terms (i.e., evidence that the perception of speech sounds can be modified by prior exposure to words and nonwords containing these sounds – cf. McQueen, Norris & Cutler, 2001).

fact that the learners were of at least intermediate level. Hence a further reason to test L2 beginners in the context of further studies. Finally, the conclusions of this study would certainly gain generalizability from testing the segmentation performances of Francophone learners of English against those of Anglophone learners of French.

*Procedural changes*. The combination of main online tasks and check-up offline tasks has helped consolidate the validity of the conclusions from this study. Yet, online psycholinguistic tasks such as word spotting always run the risk of participants consciously (hence slowly) responding to the proposed stimuli. In new tests of cue usage in L2 speech segmentation, this limitation could be overcome using an experimental technique that tests neurological activity rather than behavioral task-based responses. The technique consists in recording indicators of electrophysiological activity (known as event-related potentials, or ERPs) as participants simply listen to (or watch) stimuli. What makes the paradigm all the more suitable for testing online speech segmentation is that it collects data continuously rather than in a single measure of reaction times per trial. In addition, ERP studies of speech segmentation can easily adapt the method that consists in comparing segmentation performances between a cue-implementing condition (e.g., words with primary stress, like '*de*vil') and a control condition (e.g., words with secondary stress, like 'des*troy*'). As way of example, a pioneering study by Sanders and Neville (2003) presented Japanese ESL learners with sentence-embedded stimuli like '*de*vil' or 'des*troy*'. The learners did not show the ERPs typically associated with the acoustic correlates of syllable stress (i.e., loudness and length), suggesting that L2 listeners may use prosodic cues differently from L1 listeners. It is easy to imagine how similar ERP procedures could focus on the use of phonotactic L1 vs. L2 cues.

Connectionist simulations, finally, seem well suited to readdress the issue of cue integration in L2 segmentation. Consider the possibility that cue integration might compensate for the limited accessibility of lexical cues in the early stages of L2 acquisition (i.e., Subhypothesis 3a). Empirical research on infant speech segmentation (Redington & Chater, 1998; Christiansen & Dale, 2001) suggests that this hypothesis and the alternative scenario of 'increased integration' (i.e., Subhypothesis 3b) could be retested against each other by comparing the performances of a connectionist network before and after it has been trained to detect the relevant lexical and sublexical cues.

5.3.2. Possible new empirical avenues in the study of L2 speech segmentation

The role of attention in phonotactically-driven speech segmentation, as claimed with regard to L1 listeners (Crouzet & Bacri, 1999b) and as mentioned on several occasions in this dissertation, deserves to be further tested among L2 listeners. Such empirical work is all the more needed as SLA theories increasingly emphasize the role of attention in input-driven L2 learning (e.g., Harrington & Dennis, 2002).

Individual differences in the use of segmentation cues also call for more research. So far, related findings have been sparse and inconclusive. Thus, while Vandergrift (1997) found no significant link between gender or learning style and the preference for such or such L2 listening strategy, Harley (2000) found no evidence that age affected the ability to focus on syntax or prosody of incoming L2 utterances.

The integrability of auditory and visual cues, along with lexical and sublexical cues, is also worth exploring. If indeed, as this study suggests, cue integration may play a part in L2 speech segmentation, then it is also possible that cue integration expands beyond the encapsulated linguistic categories of lexical and sublexical cues. This assumption gains credence from multiple indications that environmental factors can influence L2 speech segmentation. Some of these influences may be negative: noise (McAllister, 1989; Bond, Moore & Gable, 1996; Levine, 2000), fast L1 speech (Conrad, 1989; Griffiths, 1990; Goh, 1997), form reduction in colloquial L1 speech (e.g., 'I hear you *wanna* quit') (Brown & Yule, 1983). But others, usually originated by L1 interlocutors, can have positive, cue-like influences: pauses (e.g., Levitt & Geoffrion, 1994), slower speech (Owen, 1996), over-segmented speech (Pica, 1994), even lip movement (Hardison, 1988; Garfunkel-Aloufy, 1992, cited in Tobin, 1997). What may need to be experimentally tested is whether such cues intervene while an L2 listener segments a target word from a stimulus L2 (or L2-sounding) input. Should preliminary findings suggest so, then more work could test the possibility of interactions between such multimodal cues and lexical/sublexical cues (as explored in the present research).

## 5.4. Pedagogical guidelines

This section develops three guidelines for reemphasizing segmentation skills in L2 listening and for initiating learners to the value of segmentation cues (notably, sublexical-phonotactic cues): 1) "diagnose" (i.e., "diagnose the students' segmentation difficulties/strategies"); 2) "encourage self-monitoring" (i.e., the students' monitoring of their own progress); 3) "encourage linguistic observation" (i.e., observation of how the L2 works, notably at the phonotactic level). The first two guidelines, though not specifically driven by the present findings, are made in an effort to promote awareness of L2 segmentation skills. The third borrows directly from the above findings.

### 5.4.1. "Diagnose"

Diagnoses can address both segmentation difficulties and segmentation strategies.

Segmentation difficulties, most apparent among beginning L2 learners, include the misuse of formulaic expressions in L2 speech, such as 'all gone' for 'gone' (Peters, 1983). Such a symptom could reveal to a teacher that a learner is continuing to perceive and memorize entire utterances as non-segmentable units.

Segmentation strategies, meanwhile, are no panacea. Indeed, L2 listeners who stystematically decode input word by word are ineffective (O'Malley, Chamot & Küpper, 1989). Yet, speech segmentation having been introduced in this thesis as key to language comprehension, one can empathize with learners willing to hone their skills at segmenting words from L2 input. A teacher who shares this concern may wish to consider the segmentation strategies sporadically reported in the L2 listening literature. These strategies, as identified by Goh (2002) using retrospective learner reports (see also Goh, 2000 and Ross, 1997), do seem susceptible of involving both lexical and sublexical cues. Here were the lexical strategies listed in Goh's reports: 'memorize words for later processing', 'visualize keyword spellings', 'focus on familiar content words'. Here were the sublexical strategies: 'anticipate aural details', 'focus on intonation features'. There was thus no indication that phonotactics can help, but the present experimental findings (as well as those of Weber and colleagues) consistently suggest otherwise.

5.4.2. "Encourage self-monitoring"

From a research standpoint, phenomenological hints that listeners are using this or that type of cue (as reported above) should be taken with caution, among other reasons because the conscious estimation of time patterns operates on a second-level scale whereas speech discrimination operates on a millisecond-level scale (Buonomano & Karmarkar, 2002). From a pedagogical standpoint, however, it should be noted that self-monitoring is one of the key metacognitive strategies of effective L2 listeners (e.g., Goh, 1997, 2002; Chan, 2005). In light of the present research, self-monitoring could consist in a learner's evaluation of how in-class phonotactic training helps them in real-life listening. An ability or, at least, a readiness for such evaluations may well help them consolidate formally acquired phonotactic skills. In any case, whether formally or informally acquired, such skills seem likely to develop through the learners' repeated observations of the relevant language structures. This point is further emphasized below.

5.4.3. "Encourage linguistic observation"

L2 teaching should not just address the communicative aspirations of L2 learners, but also emphasize the gains that learners can reap from examining the formal properties of the L2. Good language teaching, it seems, bets on the learners' metalinguistic abilities (to think about language in a decontextualized manner) as well as on their metacognitive abilities (to monitor their own progress). With this philosophy in mind, language teachers may wish to design materials and tasks that focus on the lexicon and phonotactics of the target language both explicitly and simultaneously.

Explicit learning of how L2 phonotactics works is doubly valuable. First, as mentioned earlier, it may help learners make better implicit/unconscious use of phonotactic information during online L2 speech processing. Second, "à la linguist" observations of lexical and phonotactic patterns seem affordable to beginning L2 learners with no grammatical training. It should suffice, in principle, that they notice how often or how rarely certain words and consonant clusters occur in the target language. In light of the present study, indeed, "frequency" may become a key notion for the language teacher, as much as for the psycholinguist.

Simultaneous teaching of L2 lexicon and L2 phonotactics is encouraged by two observations. First, this study suggests that lexical knowledge and phonotactic knowledge can both contribute to the gradual improvement of L2 segmentation skills. (Lexical knowledge, here, does not exclude awareness of cognacy between L1 and L2, even though cognacy has here been shown to have limited effects on L2 segmentation). Second, the phonotactic configurations of a language are most easily discovered by looking at the relevant sound arrangements inside or between words in that language. Examples of auditory tasks, in this context, could include: 1) word-to-definition matching tasks, involving the recognition of preselected words in their citation forms; 2) audio-, video or internet-based tasks involving the recognition of the same target words in reduced forms and/or embedded in sentential utterances; 3) dictation of such utterances, followed by inductive tasks aiming at discovering the permissible and less permissible conglomerations of word boundary consonants in the target language.

## 5.5. Epilogue

> *"[...] the formal task of spelling out how words are segmented from the speech stream over the time course of learning an L2 and documenting what occurs empirically is a formidable research task"*
>
> (Carroll, 2004: 250)

The "formidable research task" is in its infancy, and the present work is but a small step on the right path. Much remains to be done, also, to rehabilitate the teaching of sublexical and listening skills in the language classroom.

Nevertheless, by offering evidence that lexical and phonotactic competencies can both help language learners in segmenting L2 speech, this study will have hopefully contributed to demystify one of the most complex, vital and yet overlooked aspects of second language acquisition.

## BIBLIOGRAPHY

Altenberg, E. (2005). The perception of word boundaries in a second language. *Second Language Research* 21, 4, 325-358.

Baayen, R. H., Piepenbrock, R. and L. Gulikers (1995). *The CELEX lexical database (CD-ROM)*. University of Pennsylvania, Philadelphia, PA.

Bachman, L. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.

Bacri, N. and M. H. Banel (1995). Do metrical and phonotactic segmentation cues cooperate in spoken word recognition? In *Proceedings of the International Conference of Phonetic Sciences*, Vol. 3, pp. 608-611. Stockholm: Royal Institute of Technology and Stockholm University.

Baddeley, A. D., Thomson, N. and M. Buchanan (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior* 14, 575–589.

Balota, D. A. and J. I. Chumbley (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language* 24, 89-106.

Banel, M.-H. and N. Bacri (1997). Reconnaissance de la parole et indices de segmentation métriques et phonotactiques. *L'Année Psychologique* 97, 77–112.

Bannert, R., Nicolas, P. and M. Stridfeldt (1996). Lexical access in French: Recognizing 'identical' phrases. *Acoustical Society of America 132nd meeting*. Honolulu, HI: 2-6 December 1996.

Barbour, R. P. (1995). *Spoken Word Recognition as a Function of Lexical Knowledge and Language Proficiency Level in Adult ESL Learners*. PhD Dissertation. University of British Columbia.

Blevins, J. (2003). The independent nature of phonotactic constraints: An alternative to syllable-based approaches. In Fery, C. and R. van de Vijver (eds.), *The Syllable in Optimality Theory*. New York: Cambridge University Press.

Blevins, J. (1995). The syllable in phonological theory. In Goldsmith, J. (ed.), *the Handbook of Phonological Theory.* Cambridge, MA: Blackwell.

Bond, Z. S., Moore, T. J. and B. Gable (1996). Listening in a second language. *Proceedings of the International Conference on Spoken Language Processing.* Philadelphia, October 1996, 2510-2513.

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

Bradlow, A. R. and T. Bent (2002). The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America* 112, 1, 272-284.

Bradlow, A.R. and D. B. Pisoni (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors. *Journal of the Acoustical Society of America* 106, 2074-2085.

Broersma, M. (2002): Comprehension of non-native speech: inaccurate phoneme processing and activation of lexical competitors. *ICSLP (International Conference on Spoken Language Processing)* 261-264.

Broselow, E. (1988). Prosodic phonology and the acquisition of a second language. In Flynn, S. and. W. O'Neil (eds.), *Linguistic Theory in Second Language.* Dordrecht: Kluwer Academic Publishers.

Brown, G. and G. Yule (1983). *Discourse Analysis.* Cambridge: Cambridge University Press.

Buchwald, A. B. (2005). *Sound Structure Representation, Repair and Well-formedness: Grammar in Spoken Language Production.* PhD Dissertation. John Hopkins University.

Caramazza, A. and I. Brones (1979). Lexical access in bilinguals. *Bulletin of the Psychonomic Society* 13, 4, 212-214.

Carroll, S. E. (2002). Induction in a modular learner. *Second Language Research* 18, 224-249.

Carroll, S. E. (2004). Segmentation: Learning how to 'hear words' in the L2 speech stream. *Transactions of the Philological Society* 102, 227-254.

Champagne-Muzar, C. (1991). Le rôle des faits phonétiques dans le décodage perceptif en langue seconde : Etat de la question. *Revue québécoise de linguistique* 21, 41-55.

Chan, C.-Y. (2005). A study of metacognitive strategies in EFL Listening comprehension. *22nd International Conference on English Teaching and Learning in the Republic of China*. June 4th-5th. Taipei, Taiwan: National Taiwan Normal University.

Chen, A., Gussenhoven, C. and T. Rietveld (2004). Language-specificity in the perception of paralinguistic intonational meaning. *Language and Speech* 47, 4, 311-349.

Chesneau, G. F. (1992). *The Effects of Acoustic-Phonetic, Semantic, and Syntactic Information on Word Recognition by Elementary and Intermediate French Learners and Native French Speakers*. PhD Dissertation. Ohio State University.

Christiansen, M.H. and R.A.C. Dale (2001). Integrating distributional, prosodic and phonological information in a connectionist model of language acquisition. In *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 220-225). Mahwah, NJ: Lawrence Erlbaum.

Christie, W. M. (1977). Some multiple cues for juncture in English. *General Linguistics* 17, 212-222.

Church, K.W. (1987). *Phonological parsing in speech recognition*. Dordrecht: Kluwer Academic Publishers.

Clements, George N. (1990): The role of the sonority cycle in core syllabification. In Kingston, J. and M. E. Beckman (eds.): *Papers in Laboratory Phonology I*. Cambridge: CUP, 283-333.

Clements, G.N. and S. Jay Keyser. (1983). *CV phonology: a generative theory of the syllable*. MIT Press, Cambridge, MA.

Cohen J. D., MacWhinney, B., Flatt M. and J. Provost (1993) PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavioral Research Methods, Instruments & Computers* 25(2)257-271.

Cole, R. and J. Jakimik (1980). A model of speech perception. In R. Cole (Ed.), *Perception and Production of Fluent Speech*. Hillsdale, NJ: Erlbaum.

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology, Human Experimental Psychology* 33A, 497-505

Conrad, L. R. (1985) Semantic versus syntactic cues in listening comprehension. *Studies in Second Language Acquisition* 7, 59-72.

Conrad, L. R. (1989). The effects of time-compressed speech on native and EFL listening comprehension. *Studies in Second Language Acquisition* 11, 1, 1-16.

Content, A., Mousty, P. and M. Radeau (1990). Brulex. Une base de données lexicales informatisée pour le français écrit et parlé. *L'Année Psychologique*, *90*, 551-566.

Cooper, N., Cutler, A. and R. Wales (2002). Constraints of lexical stress on lexical access in English: Evidence from native and nonnative listeners. *Language and Speech* 45, 207-228.

Cowan, N. (1991). Recurrent speech patterns as cues to the segmentation of multisyllabic sequences. *Acta Psychologica* 11, 121-135.

Cristoffanini, P., Kirsner, K. and D. Milech (1987). Bilingual lexical representation: The status of Spanish English cognates. *The Quarterly Journal of Experimental Psychology* 38a, 367-393.

Crouzet, O. and N. Bacri (1998). Groupes consonantiques illégaux et segmentation de la parole. *Actes des 22èmes journées d'études sur la parole* (p. 275-278). Martigny, Suisse.

Crouzet, O. and N. Bacri (1999a). Phonological constraints in speech segmentation processes: Investigating levels of implementation. *Proceedings of Eurospeech '99*, pp. 2671-2674. 5th-10th September 1999, Budapest, Hungary.

Crouzet, O. and N. Bacri (1999b). Toward an attentional account of the involvement of phonotactic constraints in speech segmentation processes. *Proceedings of Forum Acusticum'99 (Joint 137th meeting of the Acoustical Society of America & 2nd convention of the European Acoustics Association)*, 14th-19th March 1999, Berlin, Germany.

Cutler, A. and S. Butterfield (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language* *31*, 218–236.

Cutler, A. and D. M. Carter (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language* 2, 133-142.

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

Cutler, A., Mehler, J., Norris, D. and J. Ségui (1989). Limits on bilingualism. *Nature* 340, 229–230.

Cutler, A., Mehler, J., Norris, D. and J. Ségui (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology* 24, 381-410.

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

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

Dahan, D., Magnuson, J. S. and M. K. Tanenhaus (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology* 42, 317-367.

Davis, M. H. (2000) *Lexical segmentation in spoken word recognition*. PhD dissertation. Birkbeck University of London.

De Bruijn, E.R.A., Dijkstra, A.F.J., Chwilla, D.J. and H. J. Schriefers (2001). Language context effects on interlingual homograph recognition: Evidence from event-related potentials and response times in semantic priming. *Bilingualism: Language and Cognition* 4 (2), 155-168.

De Groot, A. and G. Nas (1991) Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language* 30, 90-123.

Dejean de la Bâtie, B. and D. C. Bradley (1995). Resolving word boundaries in spoken French: Native and non-native strategies. *Applied Psycholinguistics* 16, 59-81.

Delattre, P. (1940). Tendances de coupe syllabique en français. *Publications of the Modern Language Association* 55, 579-595.

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

Dijkstra, A. F. J., Grainger, J. and W.J.B. Van Heuven (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language* 41, 4, 496-518.

Dijsktra, A. F. J. and J. G. Van Hell (2003). Testing the Language Mode hypothesis using trilinguals. *International Journal of Bilingual Education and Bilingualism, 6*, 2-16.

Dumay, N., Content, A. and U.H. Frauenfelder (1999). Acoustic-phonetic cues to word boundary location: evidence from word-spotting. In *Proceedings of the International Congress of Phonetic Sciences* pp. 281-284. San Francisco, CA, USA.

Dumay, N., Frauenfelder, U.H. and A. Content (2000). Acoustic-phonetic cues and lexical competition in segmentation of continuous speech. In Cutler, A., McQueen, J.M. and R. Zondervan (eds.), *Spoken Word Access Processes*. Max-Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

Dumay, N., Frauenfelder, U.H. and A. Content (2002). The role of the syllable in lexical segmentation in French: Word-spotting data. *Brain and Language* 81, 144–161.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C. and J. Mehler (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25, 1568-1578.

Eefting, W. Z. F. (1991). *Timing in talking: Tempo variation in production and its role in perception*. PhD Dissertation. Utrecht University.

Enomoto, K. (1992). Interlanguage phonology: The perceptual development of durational contrasts by English-speaking learners of Japanese. *Edinburgh Working Papers in Applied Linguistics* 2, 25-35.

Eukel, B. (1980). Phonotactic basis for word frequency effects: Implications for lexical distance metrics. *Journal of the Acoustical Society of America, 68,* s33.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In Strange, W. (ed.), *Speech Perception and Linguistic Experience*. Timonium, MD: York Press.

Flemming, E. (1995). *Auditory Representations in Phonology*. PhD Dissertation. UCLA.

Frauenfelder, U. (1985). Crosslinguistic approaches to lexical segmentation. *Linguistics* 23, 669-687.

Frenck-Mestre, C. and P. Prince (1997). Second language autonomy. *Journal of Memory and Language* 37, 481–501.

Friel, B. M. and S. M. Kennison (2001). Identifying German-English cognates, false cognates, and non-cognates: Methodological issues and descriptive norms. *Bilingualism: Language and Cognition* 4, 249-274.

Fry, D. B. (1947). *Articulation Tests*. PhD Dissertation. University of London.

Garfunkel-Aloufy, S. (1992). *Sex differences in lip-reading among native Hebrew and native English speakers (in Hebrew)*. MA Thesis. Tel-Aviv University.

Gaygen, D. and P. A. Luce (2002). Troughs and busts: Probabilistic phonotactics and lexical activation in the segmentation of spoken words in fluent speech. In McLennan, C. T., Luce, P. A., Mauner, G. and J. Charles-Luce (eds.), *University at Buffalo Working Papers on Language and Perception, Vol. 1.*

Goetry, V. and R. Kolynsky (2000). The role of rhythmic cues for speech segmentation in monolingual and bilingual listeners. *Psychological Belgica* 40, 115-152.

Goh, C. M. (1997). Metacognitive awareness and second language listeners. *ELT Journal* 51, 361-369.

Goh, C. M. (2000). A cognitive perspective on language learners' listening comprehension problems. *System* 28, 55-75.

Goh, C. M. (2002). Exploring listening comprehension tactics and their interaction patterns. *System* 30, 185-206

Golato, P. (2002). Word parsing by late-learning French-English bilinguals. *Applied Psycholinguistics* 23, 417- 446.

Goldstein, H. (1983). Word recognition in a foreign language: A study of speech perception. *Journal of Psycholinguistics Research* 12, 417-427.

Gow, D. W. and P. C. Gordon (1993). Lexical and prelexical influences on word segmentation: evidence from priming. *Journal of Experimental Psychology* 21, 344-359.

Griffiths, R. (1990). Speech rate and NNS comprehension: A preliminary study in the time-benefit analysis. *Language Learning* 40, 311-336.

Grosjean, F. (2001). The bilingual's language modes. In Nicol, J. (ed.), *One mind, Two Languages: Bilingual Language Processing.* Oxford: Blackwell.

Grossberg, S. and C. W. Myers (2000). The resonant dynamics of speech perception: Interword integration and duration-dependent backward effects. *Psychological Review* 4, 735-767.

Hall-Lew, L. (2002). *English Loanwords in Mandarin Chinese.* Unpublished Bachelor's Thesis. University of Arizona. Retrieved September 9, 2005 from www.stanford.edu/~dialect/Undergraduate%20Thesis.pdf.

Hammer, P. and G. S. Giauque (1989). *The Role of Cognates in the Teaching of French.* New York: Lang.

Hardison, D. M. (1988). Spoken word identification by native and nonnative speakers of English: Effects of training, modality, context and phonetic environment. In *Proceedings of the 5th International Conference on Spoken Language Processing.* Sydney, Australia: Causal Productions.

Harley, B. (2000). Listening strategies in ESL: Do age and L1 make a difference? *TESOL Quarterly* 34, 769-777.

Harrington, M. and S. Dennis (2002). Input-driven language learning. *Studies in Second Language Acquisition* 24, 261-268.

Henrichsen, L. E. (1984). Sandhi-variation: A filter of input for learners of ESL. *Language Learning* 34, 103-126.

Hieke, A. E. (1987). The resolution of dynamic speech in L2 listening. *Language Learning* 37, 123-140.

Hieke, A. E. (1989). Spoken language phonotactics: Implications for the ESL/EFL classroom in speech production and perception. *Language Sciences* 11, 197-213.

Hohne, E. A. and P. W. Jusczyk (1994) Two-month-old infants' sensitivity to allophonic differences. *Perception and Psychophysics* 56, 613-623.

Imai, S., Walley, A.C. and J. E. Flege (2005). Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners. *Journal of the Acoustical Society of America* 117, 896-906.

Ito, Y. (2001). Effect of reduced forms on ESL learners' input-intake process. *Second Language Studies* 20, 99-124.

Izura, C. and A. W. Ellis (2002). Age of acquisition effects in word recognition and production in first and second languages. *Psicológica* 23, 245-281.

James, A. R. (1988). *The Acquisition of Second Language Phonology: A Linguistic Theory of Developing Sound Structures*. Tubingen: Gunter Narr Verlag.

Ju, M. and P. A. Luce (2004). Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science* 15, 314-318.

Jusczyk, P. W., Cutler, A. and N. J. Redanz (1993). Infants' preference for the predominant stress patterns of English words. *Child Development* 64, 675-687.

Jusczyk, P. W., Luce, P. A. and J. Charles-Luce (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language* 33, 630–645.

Jusczyk, P. W. and R. N. Aslin (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology* 29, 1-23.

Jusczyk, P. W. and P. A. Luce (2002). Speech perception and spoken word recognition: past and present. *Ear and Hearing* 23, 2-40.

Kabak, B. (2003). *The Perceptual Processing of Second Language Consonant Clusters*. PhD Dissertation. University of Delaware.

Kacinik, N., Shears, C. and C. Chiarello (2000). Familiarity for nouns and verbs: not the same as, and better than, frequency. Paper presented at the *22nd Annual Meeting of the Cognitive Science Society, University of Pennsylvania*. August 13-15, 2000. Retrieved June 23, 2005 from http://www.ircs.upenn.edu/cogsci2000/poster1.shtml

Kellerman, E. (1983). Now you see it, now you don't. In Gass, S. and L. Selinker (eds.), *Language transfer in language learning*. Rowley, MA: Newbury House.

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

Kim, H.-Y. (1995). Intake from the speech stream: Speech elements that learners attend to. In Schmidt, R. (ed.), Attention and Awareness in Foreign Language Learning. *Technical Report*, no. 9. Second Language Teaching and Curriculum Center. University of Hawai'i Press.

Kirk, C. (2000). Syllabic cues to word segmentation. In Cutler, A., McQueen, J. M. and R. Zondervan (eds.), *Spoken Word Access Processes*. Nijmegen, The Netherlands: Max-Planck Institute for Psycholinguistics.

Kiss, G.R., Armstrong, C., Milroy, R. and J. Piper (1973). An associative thesaurus of English and its computer analysis. In Aitkin, A.J., Bailey, R.W., and N. Hamilton-Smith (eds.), *The computer and Literary Studies.* Edinburgh: University Press.

Koster, C. J. (1987). *Word Recognition in Foreign and Native Language: Effects of Context and Assimilation*. Dordrecht: Foris Publications.

Krause, M. (1988). About the perception of Russian words by German-speaking listeners. *Wissenschatfliche Zeitschrift der Friedrich Schiller Universität Jena, Gesellschafts-/Sprachwissenschafliche Reihe* 37, 2, 141-145.

Kreuz, R. J. (1987). The subjective familiarity of English homophones. *Memory and Cognition, 15*, 154-168.

Kroll, J. F. and J. S. Merves (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory & Cognition* 12, 1, 92-107.

Kroll, J. F. and E. Stewart (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 33, 149-174.

Lado, R. (1957). *Linguistics Across Cultures: Applied Linguistics for Teachers.* Ann Arbor: University of Michigan Press.

Lehiste, I. (1972). The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America,* 51, 2018–2024.

Levitt, A. G. and J. Geoffrion (1994). Prosodic characteristics of register variation in the read speech of adult native speakers of French. *ASA 128th Meeting*. Austin, Texas: 28. Nov - 2 Dec., 1994.

Levine, M. (2000). *Context Effects on the Auditory Perception of English by Adults Whose First Language is Korean*. PhD Dissertation. City University, New York.

Lipinski, J. and P. Gupta (2005). Does neighborhood density influence repetition latency for nonwords? Separating the effects of density and duration. *Journal of Memory and Language* 52, 171–192.

Liu, H., Bates, E., Powell, T. and B. Wulfeck (1997). Single word shadowing and the study of lexical access. *Applied Psycholinguistics 18, 2, 157-180.*

Luce, P.A. and D. B. Pisoni (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing* 19, 1-36.

Luce, P. A. and N. Large (2001). Phonotactics, neighborhood density, and entropy in spoken word recognition. *Language and Cognitive Processes* 16, 565-581.

McAllister, R. (1989). Studies of spoken language comprehension in L2 speakers. *Speech Transmission Laboratory Quarterly Progress and Status Report* 1, 59-62.

McClelland, J. L. and J. L. Elman (1986). The TRACE model of speech perception. *Cognitive Psychology* 18, 1-86.

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

McQueen, J. M., Norris, D. G. and A. Cutler (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20,* 621–638.

McQueen, J. M., Norris, D. G. and A. Cutler (2001). Can lexical knowledge modulate prelexical representations over time?. In Smits, R., Kingston, J., Nearey, T.M. and R. Zondervan (eds.), *Proceedings of SPRAAC (Workshop on Speech Recognition as Pattern Classification*. Nijmegen: MPI Psycholinguistics.

Major, R. (2001). *Foreign Accent*. Mahwah, NJ : Lawrence Erlbaum.

Major, R.C. and K. Eunyi (1999). The similarity differential rate hypothesis. *Language Learning* 49, 151-183.

Marian, V. and M. Spivey (1999). Activation of Russian and English cohorts during bilingual spoken word recognition. In Hahn, M. and Stoness, S. C. (eds.). *Proceedings of the Twenty-first Annual Conference of the Cognitive Science Society*. Mahwah, New Jersey: Lawrence Erlbaum.

Marian, V. and M. Spivey (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition* 6, 2, 2003, 97–115.

Marian, V., Spivey, M and J. Hirsch (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. *Brain and Language* 86, 70–82.

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

Marslen-Wilson, W. D. (1990). Activation, competition, and frequency in lexical access. In Altmann, G.T.M. (ed.), *Cognitive models of speech processing*. Cambridge, MA: MIT Press.

Matter, J. F. (1986). *A la recherche des frontières perdues: Etude sur la perception de la parole en français*. PhD thesis, University of Utrecht. Amsterdam: De Werelt.

Mattys, S.L. and P. W. Jusczyk (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition* 78, 91-121.

Mattys, S. L., Jusczyk, P. W., Luce, P. A. and J. L. Morgan (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology* 38, 465-494.

Mayo, L. H., Florentine, M. and S. Buus (1997) Age of second-language acquisition and perception of speech in noise. *Journal of Speech, Language and Hearing* 40, 3, 686-693.

Meador, D., Flege, J. and I.R.A. MacKay (2000). Factors affecting the recognition of words in a second language. *Bilingualism: Language and Cognition* 3, 55-67.

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

Meyer, D.E. and R. W. Schvaneveldt (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology* 90, 227-234.

Moudraia, O. (2001). Lexical Approach to Second Language Teaching. *ERIC Clearinghouse on Languages and Linguistics* June 1001. Retrieved November 7, 2005 from http://www.cal.org/ericcll/digest/0102lexical.html

Nabelek, A.K. and A. M. Donahue (1984). Perception of consonants in reverberation by native and non-native listeners. *Journal of the Acoustical Society of America* 75, 632-634.

Nakatani, L. (unpublished), *Fission and Fusion Processes in Speech Perception.* Unpublished manuscript. Murray Hill, NJ: Bell Laboratories.

Nakatani, L. and K. Dukes (1977). Locus of segmental cues for word juncture. *Journal of the Acoustical Society of America* 62 (3), 714-719.

Nooteboom, S. G. and P. J. M. Truin (1980). The recognition of fragments of spoken words by native and non-native listeners. *IPO Annual Progress Report* 15, 42-47.

Norris, D. G. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition* 52, 189-234.

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

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

Odlin, T. (1989). *Language Transfer*. Cambridge: Cambridge University Press.

Oller, J. and S. Ziahosseiny (1970). The contrastive analysis hypothesis and spelling errors'. *Language Learning* 20, 183-89

O'Malley, J. M., Chamot, A. U. and L. Kupper (1989). Listening comprehension strategies in second language acquisition. *Applied Linguistics* 10, 418-437.

Otake, T., Hatano, G. and K. Yoneyama (1996). Speech segmentation by Japanese listeners. In T. Otake and A. Cutler (eds.), *Phonological Structure and Language Processing: Cross-Linguistic Studies*. Berlin: Mouton / De Gruyter.

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

Owen, D. (1996). Do teachers modify their speech according to the proficiency of their students ? *English Language Teacher Education and Development* 2, 1-31.

Pál, A. (2000). *The Role of Cross-Linguistic Formal Similarity in Hungarian-German Bilingual Learners of English as a Foreign Language*. PhD dissertation. Universät Potsdam.

Parker, S. G. (2002). *Quantifying the Sonority Hierarchy*. PhD dissertation. University of Massachusetts, Amherst.

Pearman, A. (2004). Sound change in the learner: The perception of connected speech. University of Illionois at Urbana-Champaign. Urbana-Champaign, IL, 24-26 June, 2004: *9th Conference on Laboratory Phonology.*

Peereman, R. and A. Content (1999). LexOP: A lexical database with Orthography-Phonology statistics for French monosyllabic words. *Behavior Research Methods, Instruments, & Computers* 31, 376-379

Peters, A. (1983). *The Units of Language Acquisition*. Cambridge: Cambridge University Press.

Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes and outcomes? *Language Learning* 44, 493-527.

Pilon, R. (1981). Segmentation of speech in a foreign language. *Journal of Psycholinguistics Research* 10, 113-122.

Poulisse, N. and T. Bongaerts (1994). First language use in second language production. *Applied Linguistics* 15, 36-57.

Prince, A. and P. Smolensky (1993). *Optimality theory: Constraint Interaction in Generative Grammar*. Technical Report 2, Rutgers University Center for Cognitive Science.

Quené, H. (1992). Durational cues for word segmentation in Dutch. *Journal of Phonetics* 20, 331-350.

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

Radeau M., Morais J., Mousty P. and P. Bertelson (2000). The effect of speaking rate on the role of the uniqueness point in spoken word recognition, *Journal of Memory and Language* 42, 406-422.

Rastle, K., Harrington, J. and M. Coltheart (2002). 358534 nonwords: The ARC nonword database. *Quarterly Journal of Experimental Psychology* 55A, 1339-1362.

Redington, M. and N. Chater (1998). Connectionist and statistical approaches to language acquisition: A distributional perspective. *Language and Cognitive Processes* 13, 129-191.

Rialland, A. (1994). The phonology and phonetics of extrasyllabicity in French. In Keating, P. A. (ed.), *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form,*. New York: Cambridge University Press.

Robert, P. (1986). *Micro-Robert. Dictionnaire du Français primordial.* Paris: Dictionnaires Le Robert.

Rosa, M. (2002). Don'cha know? A survey of ESL teachers' perspectives on reduced forms instruction. *Second Language Studies* 21, 49-78.

Ross, S. (1997). An introspective analysis of listener inferencing on a second language listening test. *Applied Linguistics* 18, 3-19.

Rost, M. (1990) *Listening in Language Learning*. London, Longman.

Rubenstein, H., Garfield, L. and J. A. Milikan (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior* 18, 757-767.

Sanders, L. D. and H. J. Neville (2000). Lexical, syntactic, and stress-pattern cues for speech segmentation. *Journal of Speech, Language, and Hearing Research* 43, 1301-1321.

Sanders, L. D., Neville, H. J. and M. G. Woldorff (2002). Speech segmentation by native and non-native speakers: The use of lexical, syntactic, and stress-pattern cues. *Journal of Speech, Language and Hearing Research* 45, 519-530.

Sanders, L. D. and H. J. Neville (2003). An ERP study of continuous speech processing. II. Segmentation, semantics, and syntax in non-native speakers. *Cognitive Brain Research* 15, 228-240.

Sebastián-Gallés, N., Echeverría, S. and L. Bosch (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language* 52, 240–255.

Selkirk, E. O. (1984). On the major class features and syllable theory. In Aronoff, M. and R. T. Oerhle (eds). *Language Sound Structure: Studies in phonology presented to Morris Halle by his teacher and students.* Cambridge, Mass.: MIT Press. 107-136.

Sinai, A. and H. Pratt (2002). Electrophysiological Evidence for Priming in Response to Words and Pseudowords in First and Second Language. *Brain and Language* 80, 2, 240-252.

Sinor, M. (2002). Durational cues in the identification of word onsets by L2 learners of French. Poster presented at the *Third International Conference on the Mental Lexicon*. Banff, Alberta. 6. October 2002.

Sinor, M. (2003). Temporal comparison of word boundary consonants in French: The speaker-specificity issue. *The LACUS Forum* 29, 315-325.

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

Sobkowiak (unpublished). *Rule-based and Empirical Rating of Perceived Phonetic Difficulty of English Words According to Polish Learners: Does Frequency Matter?* Retrieved June 29, 2005 from http://elex.amu.edu.pl/ifa/plm/ abstracts/plm_2004_abs_ sobkowiak _ 2.htm

Spivey, M. and V. Marian (1999). Cross talk between native and second languages: partial activation of an irrelevant lexicon. *Psychological Science* 10, 181-84.

Swan, M. and B. Smith (2001). *Learner English: A Teacher's Guide to Interference and other Problems.* 2nd edition. Cambridge: Cambridge University Press.

Taft, M. and G. Hambly (1986). Exploring the cohort model of spoken word recognition. *Cognition 22*, 259–282.

Talamas, A., Kroll, J. F. and R. DuFour (1999). From form to meaning: Stages in the acquisition of second-language vocabulary. *Bilingualism: Language and Cognition* 2, 1, 45–58.

Takayanagi, S., Dirks, D. D. and A. Moshfegh (2002). Lexical and talker effects on word recognition among native and non-native listeners with normal and impaired hearing. *Journal of Speech Language and Hearing Research* 45, 585–597.

Tarone, E. (1974). Speech perception in second language acquisition: A suggested model. *Language Learning* 24, 223-233.

Tobin, Y. (1997). *Phonology as Human Behaviour: Theoretical Implications and Clinical Applications.* Duke University Press.

Treiman, R., Kessler, B., Knewasser, S., Tincoff, R. and M. Bowman (2000). English speakers' sensitivity to phonotactic patterns. In Broe, M. B. and J. B. Pierrehumbert (eds.), *Papers in Laboratory Phonology V: Acquisition and the lexicon.* Cambridge, England: Cambridge University Press.

Tyler, L. K., Voice, J. K. and H. E. Moss (2000). The interaction of meaning and sound in spoken word recognition. *Psychonomic Bulletin and Review 7*, 320–326.

van Heuven, W. J. B., Dijkstra, T. and J. Grainger (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language* 39, 458-483.

van Wijngaarden, S. J. (2001). Intelligibility of native and non-native Dutch speech. *Speech Communication* 35, 103–113.

van Wijngaarden, S. J., Steeneken H. J. and T. Hougast (2002). Quantifying the intelligibility of speech in noise for non-native listeners. *Journal of the Acoustical Society of America* 111, 4, 1906-1916.

Vitevitch, M. S. and P. A. Luce (1998). When words compete: Levels of processing in spoken word perception. *Psychological Science* 9, 325–329.

Vitevitch, M. S. and P. A. Luce (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language 40,* 374-408.

Vitevitch, M.S., Luce, P.A., Charles-Luce, J. and D. Kemmerer (1996). Phonotactic and metrical influences on the perception of spoken and nonsense words. *Language and Speech 40,* 47-62.

Vitevitch, M. S., Luce, P. A., Charles-Luce, J. and D. Kemmerer (1997). Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech* 40, 47-62.

Vitevitch, M. S. and P. A. Luce (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language* 52, 193-204.

Vogel, D. and H. Winitz (1989). Word isolation and meaning in segmentation. *Journal of Psycholinguistic Research* 18, 473-484.

Wakefield, J. A., Doughtie, E. G. and B.-H. Lee Yom (1974). The identification of structural components of an unknown language. *Journal of Psycholinguistics Research* 3, 261-269.

Weber, A. (2000). The role of phonotactics in the segmentation of native and non-native continuous speech. In A. Cutler, J. McQueen and R. Zondervan (eds.), *Proceedings of the Workshop on Spoken Access Processes.* Nijmegen, the Netherlands: Max Planck Institute for Psycholinguistics.

Weber, A. (2000). The role of phonotactics in the segmentation of native and non-native continuous speech. In A. Cutler, J. McQueen, & R. Zondervan (eds.), *Proceedings of the Workshop on Spoken Access Processes,* (pp. 143-146). Nijmegen, the Netherlands: Max Planck Institute for Psycholinguistics.

Weber, A. and A. Cutler (2004). Lexical competition in non-native spoken- word recognition. *Journal of Memory and Language* 50, 1-25.

Weber, A. and A. Cutler (2006). First-language phonotactics in second-language listening. *Journal of the Acoustical Society of America.*

Wioland, F. (1985). *Les structures syllabiques du français: fréquence et distribution des phonèmes consonantiques; contraintes idiomatiques dans les séquences consonantiques*. Genève: Slatkine, collection Travaux de Linguistique quantitative.

Woutersen, M., Cox, A., Weltens, B. and K. de Bot (1994). Lexical aspects of standard dialect bilingualism. *Applied Psycholinguistics* 15, 447–473.

Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort*. Cambridge: Addison-Wesley.

## Appendix A. Distribution of CC clusters in French and English

*Onsets*

| Type of consonant combination | Cluster item | Occurrence frequency in English (%) | Occurrence frequency in French (%) |
|---|---|---|---|
| Obstruent-Liquid (OBLI) | [bl] | 4.0 | 4.0 |
| | [fl] | 5.1 | 5.4 |
| | [gl] | 3.0 | 3.7 |
| | [kl] | 6.2 | 5.7 |
| | [pl] | 3.8 | 6.0 |
| | [br, bʁ] | 8.3 | 9.9 |
| | [fr, fʁ] | 3.7 | 8.5 |
| | [gr, gʁ] | 7.8 | 9.1 |
| | [kr, kʁ] | 6.3 | 10.5 |
| | [pr, pʁ] | 4.4 | 8.5 |
| | [tr, tʁ] | 5.7 | 10.8 |
| Obstruent-Obstruent (OBOB) | [ps] | 0 | 0.6 |
| | [sf] | 0.2 | 0.6 |
| | [sk] | 4.8 | 2.3 |
| | [sm] | 3.8 | 0.6 |
| | [sn] | 3.4 | 0.6 |
| | [sp] | 6.5 | 1.1 |
| | [st] | 8.5 | 5.7 |

*Appendix A continues next page*

*Offsets*

| Type of consonant combination | Cluster item | Occurrence frequency in English (%) | in French (%) |
|---|---|---|---|
| Obstruent-Obstruent (OBOB) | [dz] | 0.7 | 0 |
| | [ft] | 3.5 | 0.3 |
| | [gd] | 0.5 | 0 |
| | [gm] | 0.3 | 0.2 |
| | [ks] | 0.5 | 1.4 |
| | [kt] | 0.9 | 1.4 |
| | [mf] | 0.5 | 0 |
| | [mk] | 0.1 | 0 |
| | [mt] | 0.1 | 0 |
| | [mp] | 5.3 | 0 |
| | [nd] | 7.0 | 0 |
| | [nf] | 0.1 | 0 |
| | [nk] | 5.8 | 0 |
| | [ns] | 5.5 | 0 |
| | [nt] | 11.6 | 0 |
| | [pf] | 0.1 | 0 |
| | [ps] | 0.5 | 1.0 |
| | [pt] | 0.9 | 0.7 |
| | [sk] | 2.3 | 3.1 |
| | [sm] | 0 | 0.3 |
| | [sp] | 1.1 | 0.3 |
| | [st] | 14.0 | 2.8 |
| Liquid-Obstruent (LIOB) | [lb] | 0.7 | 0.2 |
| | [ld] | 4.3 | 0.7 |
| | [lf] | 1.4 | 0.7 |
| | [lk] | 1.8 | 1.4 |
| | [lm] | 0.9 | 1.0 |
| | [ln] | 0.2 | 0.3 |
| | [lp] | 0.9 | 1.0 |
| | [ls] | 0.2 | 0.3 |
| | [lt] | 7.0 | 2.4 |
| | [lv] | 1.1 | 0 |

**Appendix B. Stimuli used for the pilot study of phonotactic cues (Exp. 1 and 2)**

*LEG-ILLEG (legal offset – illegal onset)*

bimpnon, bintmot, bonssef, dinkkep, fumpsaf, gontfas, gostbup, kestpab, kunkfip, lunkmin, menktes, nantlis, niltmaf, nintbes, nistkem, nulttaf, poftteb, pustfes, santtib, seltkib, sesttol, simppef, tenkbam

*ILLEG-LEG (illegal offset – legal onset)*

babflim, bofklud, bonplak, fakblep, fapklem, fetkluf, fikflaf, fimblat, fotplif, fufploz, fupplef, kukplod, latblon, lenblap, lomklon, lutflet, nebbluk, nefflun, nepflon, nosblem, nukklag, safbleg, tibklug

*LEG-LEG (legal offset – legal onset)*

bamplek, bonspef, domplaf, femplon, fonstig, gamplok, kensmuk, lenklaf, linsnaf, linstof, minskos, ponklum, venklof

*ILLEG-ILLEG (illegal offset – illegal onset)*

femtkas, fotfsut, mekpfap, minfpom, nimkfon, nootkpum, papfken, pekfmip, tetpsop

## Appendix C. Responses to printed legal English CC clusters (Exp. 1)

### Responses to legal onset clusters

| Onset cluster | Stimulus String | Stimulus Category | Frequency of string-medial n-gram | | | Transcription Type | | Onset Preservation Rate (%) |
|---|---|---|---|---|---|---|---|---|
| | | | trigram | offset bigram | onset bigram | CC-C | C-CC | |
| [bl] | safbleg | ILLEG-LEG | 0 | 11 | 7022 | 1 | 39 | 97.5 |
| | fimblat | ILLEG-LEG | 225 | 2103 | 7022 | 8 | 32 | 80.0 |
| | fakblep | ILLEG-LEG | 0 | 30 | 7022 | 0 | 40 | 100.0 |
| | nosblem | ILLEG-LEG | 0 | 202 | 7022 | 1 | 39 | 97.5 |
| | lenblap | ILLEG-LEG | 0 | 70 | 7022 | 1 | 39 | 97.5 |
| | latblon | ILLEG-LEG | 0 | 93 | 7022 | 0 | 40 | 100.0 |
| | nebbluk | ILLEG-LEG | 66 | 414 | 7022 | 2 | 38 | 95.0 |
| [fl] | babflim | ILLEG-LEG | 0 | 0 | 1894 | 0 | 40 | 100.0 |
| | nepflon | ILLEG-LEG | 0 | 31 | 1894 | 3 | 37 | 92.5 |
| | fikflaf | ILLEG-LEG | 0 | 75 | 1894 | 0 | 40 | 100.0 |
| | lutflet | ILLEG-LEG | 1 | 179 | 1894 | 1 | 39 | 97.5 |
| | nefflun | ILLEG-LEG | 57 | 4086 | 1894 | 6 | 34 | 85.0 |
| [kl] | venklof | LEG-LEG | 48 | 1751 | 28048 | 8 | 32 | 80.0 |
| | lenklaf | LEG-LEG | 0 | 1751 | 260 | 5 | 35 | 87.5 |
| | ponklum | LEG-LEG | 3 | 2120 | 1894 | 3 | 37 | 92.5 |
| | lomklon | ILLEG-LEG | 0 | 2 | 436 | 12 | 28 | 70.0 |
| | fetkluf | ILLEG-LEG | 0 | 1 | 436 | 0 | 40 | 100.0 |
| | tibklug | ILLEG-LEG | 0 | 0 | 436 | 0 | 40 | 100.0 |
| | fapklem | ILLEG-LEG | 0 | 15 | 436 | 0 | 40 | 100.0 |
| | bofklud | ILLEG-LEG | 0 | 0 | 436 | 3 | 37 | 92.5 |
| | nukklag | ILLEG-LEG | 0 | 5 | 436 | 4 | 36 | 90.0 |
| [pl] | fufploz | ILLEG-LEG | 0 | 1 | 7032 | 0 | 40 | 100.0 |
| | bonplak | ILLEG-LEG | 31 | 80 | 7032 | 0 | 40 | 100.0 |
| | kukplod | ILLEG-LEG | 0 | 12 | 7032 | 0 | 40 | 100.0 |
| | fotplif | ILLEG-LEG | 5 | 52 | 7032 | 3 | 37 | 92.5 |
| | fupplef | ILLEG-LEG | 334 | 2887 | 7032 | 8 | 32 | 80.0 |
| | gamplok | LEG-LEG | 0 | 1751 | 1969 | 5 | 35 | 87.5 |
| | domplaf | LEG-LEG | 1 | 1751 | 1230 | 8 | 32 | 80.0 |
| | femplon | LEG-LEG | 0 | 141 | 1894 | 4 | 36 | 90.0 |
| | bamplek | LEG-LEG | 30 | 517 | 7032 | 3 | 37 | 92.5 |
| [sk] | minskos | LEG-LEG | 0 | 4043 | 1230 | 4 | 36 | 90.0 |
| [sm] | kensmuk | LEG-LEG | 0 | 4043 | 1969 | 8 | 32 | 80.0 |
| [sn] | linsnaf | LEG-LEG | 0 | 4043 | 260 | 6 | 34 | 85.0 |
| [sp] | bonspef | LEG-LEG | 0 | 4043 | 4395 | 3 | 37 | 92.5 |
| [st] | fonstig | LEG-LEG | 34 | 4043 | 28048 | 4 | 36 | 90.0 |
| | linstof | LEG-LEG | 34 | 4043 | 28048 | 7 | 33 | 82.5 |

*Appendix C continues next page*

## Responses to legal offsets

| Offset cluster | Stimulus String | Stimulus Category | Frequency of string-medial n-gram | | | Transcription Type | | Offset Preservation Rate (%) |
|---|---|---|---|---|---|---|---|---|
| | | | trigram | offset bigram | onset bigram | CC-C | C-CC | |
| [ft] | poftteb | LEG-ILLEG | 0 | 2842 | 5706 | 40 | 0 | 100.0 |
| [lt] | seltkib | LEG-ILLEG | 0 | 2697 | 1 | 40 | 0 | 100.0 |
| | niltmaf | LEG-ILLEG | 0 | 2697 | 669 | 40 | 0 | 100.0 |
| | nulttaf | LEG-ILLEG | 0 | 2697 | 5706 | 40 | 0 | 100.0 |
| [mp] | gamplok | LEG-LEG | 0 | 1751 | 1969 | 5 | 35 | 12.5 |
| | domplaf | LEG-LEG | 1 | 1751 | 1230 | 8 | 32 | 20.0 |
| | femplon | LEG-LEG | 0 | 141 | 1894 | 4 | 36 | 10.0 |
| | bamplek | LEG-LEG | 30 | 517 | 7032 | 3 | 37 | 7.5 |
| | fumpsaf | LEG-ILLEG | 115 | 5167 | 1751 | 40 | 0 | 100.0 |
| | bimpnon | LEG-ILLEG | 1 | 5167 | 21 | 40 | 0 | 100.0 |
| | simppef | LEG-ILLEG | 2 | 5167 | 2887 | 40 | 0 | 100.0 |
| [nk] | venklof | LEG-LEG | 48 | 1751 | 28048 | 8 | 32 | 20.0 |
| | lenklaf | LEG-LEG | 0 | 1751 | 260 | 5 | 35 | 12.5 |
| | ponklum | LEG-LEG | 3 | 2120 | 1894 | 3 | 37 | 7.5 |
| | tenkbam | LEG-ILLEG | 0 | 1712 | 30 | 40 | 0 | 100.0 |
| | lunkmin | LEG-ILLEG | 5 | 1712 | 31 | 40 | 0 | 100.0 |
| | menktes | LEG-ILLEG | 1 | 1712 | 24 | 39 | 1 | 97.5 |
| | kunkfip | LEG-ILLEG | 14 | 1712 | 75 | 40 | 0 | 100.0 |
| | dinkkep | LEG-ILLEG | 0 | 1712 | 5 | 39 | 1 | 97.5 |
| [ns] | minskos | LEG-LEG | 0 | 4043 | 1230 | 4 | 36 | 10.0 |
| | kensmuk | LEG-LEG | 0 | 4043 | 1969 | 8 | 32 | 20.0 |
| | linsnaf | LEG-LEG | 0 | 4043 | 260 | 6 | 34 | 15.0 |
| | bonspef | LEG-LEG | 0 | 4043 | 4395 | 3 | 37 | 7.5 |
| | fonstig | LEG-LEG | 34 | 4043 | 28048 | 3 | 37 | 7.5 |
| | linstof | LEG-LEG | 50 | 10349 | 261 | 7 | 33 | 17.5 |
| | bonssef | LEG-ILLEG | 0 | 10349 | 9802 | 4 | 36 | 92.5 |
| [nt] | nintbes | LEG-ILLEG | 1 | 24522 | 93 | 40 | 0 | 100.0 |
| | gontfas | LEG-ILLEG | 9 | 24522 | 179 | 39 | 1 | 97.5 |
| | nantlis | LEG-ILLEG | 927 | 24522 | 3731 | 37 | 3 | 92.5 |
| | bintmot | LEG-ILLEG | 61 | 24522 | 669 | 40 | 0 | 100.0 |
| | santtib | LEG-ILLEG | 0 | 24522 | 5706 | 40 | 0 | 100.0 |
| [st] | kestpab | LEG-ILLEG | 7 | 28048 | 52 | 40 | 0 | 100.0 |
| | pustfes | LEG-ILLEG | 12 | 28048 | 179 | 40 | 0 | 100.0 |
| | nistkem | LEG-ILLEG | 0 | 28048 | 1 | 40 | 0 | 100.0 |
| | gostbup | LEG-ILLEG | 6 | 28048 | 93 | 40 | 0 | 100.0 |
| | sesttol | LEG-ILLEG | 0 | 28048 | 5706 | 39 | 1 | 97.5 |

## Appendix D. Responses to spoken legal English CC clusters (Exp. 2)

### Responses to legal onset clusters

| Onset cluster | Stimulus String | Stimulus Category | Transcription Type | | | Onset Preservation Rate (%) |
|---|---|---|---|---|---|---|
| | | | CC-C | C-CC | Other | |
| [bl] | safbleg | ILLEG-LEG | 0 | 49 | 0 | 100.0 |
| | fimblat | ILLEG-LEG | 2 | 44 | 3 | 89.8 |
| | fakblep | ILLEG-LEG | 0 | 42 | 7 | 85.7 |
| | nosblem | ILLEG-LEG | 2 | 40 | 7 | 81.6 |
| | lenblap | ILLEG-LEG | 0 | 36 | 13 | 73.5 |
| | latblon | ILLEG-LEG | 0 | 29 | 20 | 59.1 |
| | nebbluk | ILLEG-LEG | 0 | 9 | 40 | 18.4 |
| [fl] | babflim | ILLEG-LEG | 0 | 49 | 0 | 100 |
| | nepflon | ILLEG-LEG | 1 | 48 | 0 | 97.9 |
| | fikflaf | ILLEG-LEG | 0 | 47 | 2 | 95.9 |
| | lutflet | ILLEG-LEG | 0 | 42 | 7 | 85.7 |
| | nefflun | ILLEG-LEG | 5 | 18 | 26 | 36.7 |
| [kl] | venklof | LEG-LEG | 7 | 32 | 10 | 65.3 |
| | lenklaf | LEG-LEG | 5 | 42 | 2 | 85.7 |
| | ponklum | LEG-LEG | 9 | 36 | 4 | 73.5 |
| | lomklon | ILLEG-LEG | 3 | 41 | 5 | 83.7 |
| | fetkluf | ILLEG-LEG | 0 | 39 | 10 | 79.6 |
| | tibklug | ILLEG-LEG | 0 | 39 | 10 | 79.6 |
| | fapklem | ILLEG-LEG | 0 | 26 | 23 | 53.0 |
| | bofklud | ILLEG-LEG | 0 | 12 | 37 | 24.5 |
| | nukklag | ILLEG-LEG | 0 | 20 | 29 | 40.8 |
| [pl] | fufploz | ILLEG-LEG | 1 | 47 | 1 | 95.9 |
| | bonplak | ILLEG-LEG | 0 | 43 | 6 | 87.8 |
| | kukplod | ILLEG-LEG | 0 | 42 | 7 | 85.7 |
| | fotplif | ILLEG-LEG | 0 | 19 | 30 | 38.8 |
| | fupplef | ILLEG-LEG | 0 | 45 | 4 | 91.8 |
| | gamplok | LEG-LEG | 15 | 23 | 11 | 46.9 |
| | domplaf | LEG-LEG | 20 | 25 | 4 | 51.0 |
| | femplon | LEG-LEG | 0 | 41 | 8 | 83.7 |
| | bamplek | LEG-LEG | 0 | 29 | 20 | 59.2 |
| [sk] | minskos | LEG-LEG | 24 | 18 | 7 | 36.7 |
| [sm] | kensmuk | LEG-LEG | 17 | 30 | 2 | 61.2 |
| [sn] | linsnaf | LEG-LEG | 9 | 38 | 2 | 77.6 |
| [sp] | bonspef | LEG-LEG | 14 | 20 | 15 | 40.8 |
| [st] | fonstig | LEG-LEG | 8 | 25 | 16 | 51.0 |
| | linstof | LEG-LEG | 12 | 32 | 5 | 65.3 |

*Appendix D continues next page*

**Responses to legal offset clusters**

| Offset Cluster | Stimulus String | Stimulus Category | Transcription Type: CC-C | C-CC | Other | Onset Preservation Rate (%) |
|---|---|---|---|---|---|---|
| [lt] | seltkib | LEG-ILLEG | 1 | 0 | 48 | 3.5 |
| | niltmaf | LEG-ILLEG | 11 | 0 | 38 | 22.5 |
| | nulttaf | LEG-ILLEG | 5 | 0 | 44 | 11.4 |
| [mp] | gamplok | LEG-LEG | 15 | 23 | 11 | 30.6 |
| | domplaf | LEG-LEG | 20 | 25 | 4 | 40.8 |
| | femplon | LEG-LEG | 0 | 41 | 8 | 0.0 |
| | bamplek | LEG-LEG | 0 | 29 | 20 | 0.0 |
| | fumpsaf | LEG-ILLEG | 1 | 0 | 48 | 3.5 |
| | bimpnon | LEG-ILLEG | 2 | 0 | 47 | 4 |
| | simppef | LEG-ILLEG | 5 | 0 | 44 | 11.4 |
| [nk] | venklof | LEG-LEG | 7 | 32 | 10 | 14.3 |
| | lenklaf | LEG-LEG | 5 | 42 | 2 | 10.2 |
| | ponklum | LEG-LEG | 9 | 36 | 4 | 18.4 |
| | tenkbam | LEG-ILLEG | 41 | 0 | 8 | 83.7 |
| | lunkmin | LEG-ILLEG | 29 | 0 | 20 | 59.2 |
| | menktes | LEG-ILLEG | 25 | 0 | 24 | 51 |
| | kunkfip | LEG-ILLEG | 8 | 0 | 41 | 16.3 |
| | dinkkep | LEG-ILLEG | 21 | 0 | 28 | 42.9 |
| [ns] | minskos | LEG-LEG | 24 | 18 | 7 | 49.0% |
| | kensmuk | LEG-LEG | 17 | 30 | 2 | 34.7% |
| | linsnaf | LEG-LEG | 9 | 38 | 2 | 18.4% |
| | bonspef | LEG-LEG | 14 | 20 | 15 | 28.6% |
| | fonstig | LEG-LEG | 8 | 25 | 16 | 16.3% |
| | linstof | LEG-LEG | 12 | 32 | 5 | 24.5% |
| | bonssef | LEG-ILLEG | 4 | 0 | 45 | 8.2 |
| [nt] | nintbes | LEG-ILLEG | 33 | 0 | 16 | 67.3 |
| | gontfas | LEG-ILLEG | 6 | 0 | 43 | 12.2 |
| | nantlis | LEG-ILLEG | 13 | 0 | 36 | 26.5 |
| | bintmot | LEG-ILLEG | 18 | 0 | 31 | 36.7 |
| | santtib | LEG-ILLEG | 10 | 0 | 39 | 20.4 |
| [st] | kestpab | LEG-ILLEG | 30 | 0 | 19 | 61.2 |
| | pustfes | LEG-ILLEG | 24 | 0 | 25 | 49.0 |
| | nistkem | LEG-ILLEG | 26 | 0 | 23 | 53.1 |
| | gostbup | LEG-ILLEG | 7 | 0 | 42 | 14.3 |
| | poftteb | LEG-ILLEG | 0 | 0 | 49 | 0.0 |
| | sesttol | LEG-ILLEG | 4 | 3 | 42 | 8.2 |

## Appendix E - Language Background Questionnaire

Biographical Data

Sex: M ❐ F ❐ Age: ___ Occupation: Student ❐ Other ❐ Affiliation: ___________

Language background check

First language (L1): My first language is: French ❐ Chinese ❐
Is French your L1 alongside another language? Yes ❐ No ❐
If yes, which language (e.g., Lingala, Kirundi)? ___________________

ESL Experience

Are you enrolled in an ESL course? Yes ❐ No ❐ If yes, at what level? ________________

Are you taking Anglophone university courses? Yes ❐ No ❐ If yes, at what level? __________

How long have you been studying English? ____________________________________

Where did you start learning English? primary school ❐ secondary school ❐ university ❐

In what context are you using English? class only ❐ work only ❐ socially only ❐
socially, sometimes ❐ socially, often ❐

Language modes

| | | | |
|---|---|---|---|
| How often do you use your L1 (French or Chinese)? | always ❐ | often ❐ | rarely ❐ |
| How often do you use your L2 (English)? | always ❐ | often ❐ | rarely ❐ |
| How often do you use both your L1 and your L2? | always ❐ | often ❐ | rarely ❐ |

Language Stability/Proficiency

| | | | |
|---|---|---|---|
| At which level would you rate your English? | beginner ❐ | advanced ❐ | fluent ❐ |
| How well do you *speak* in English? | very well ❐ | ok ❐ | badly ❐ |
| How well do you *write* in English? | very well ❐ | ok ❐ | badly ❐ |
| How well do you *read* in English? | very well ❐ | ok ❐ | badly ❐ |
| How well do you *listen* in English? | very well ❐ | ok ❐ | badly ❐ |

## Appendix F – Quantitative and qualitative assessment of L2 proficiency

| L2 Proficiency group | Subject number | Online scores (word-spotting, lexical decision – average) | Current professional / educational status | Prior ESL training | Age | Additional information |
|---|---|---|---|---|---|---|
| Intermediate | Chantal | 36, 54 – 45 | ESLP | ESL-3 | 35 | L2A, AL |
| | Manon | 44, 60 – 52 | CEGEP | SEC-BIT-1-2 | 20 | L2P, PL |
| | Marcel | 40, 67 – 54 | CEGEP | SEC-BIT-1-2 | 19 | L2P, AL |
| | Veronique | 40, 68 – 54 | CEGEP | SEC-BIT-1-2 | 18 | L2P, PL |
| | Marvin | 41, 66 – 54 | CEGEP | SEC-BIT-1-2 | 19 | L2P, PL |
| | Gabrielle | 43, 65 – 54 | CEGEP | SEC-BIT-1-2 | 19 | L2P, PL |
| | Magali | 45, 62 – 54 | CEGEP | SEC-BIT-1-2 | 18 | L2P, GL |
| | Sophie 1 | 47, 67 – 57 | CEGEP | SEC-BIT-1-2 | 20 | L2P, PL |
| | Floriane | 47, 69 – 58 | ESLU-Beg | SEC | 25 | CL1 Kirundi, L2P, PL |
| | Dominic | 49, 66 – 58 | CEGEP | SEC-BIT-2 | 17 | L2P, AL |
| | Marc | 49, 67 – 58 | MIL | SEC-BIT-14 | 33 | L2A, AL |
| | Laurie | 51, 67 – 59 | CEGEP | SEC-BIT-2 | 19 | L2P, PL |
| | Anne | 51, 68 – 60 | RA | SEC | 27 | L2A, PL |
| | Landry | 53, 68 – 61 | ESLU-Beg | SEC | 19 | CL1 Kirundi, L2A, AL |
| | Omayra | 60, 61 – 61 | ESLU-Beg | SEC | 19 | CL1 Hausa, L2P, AL |
| | Billy | 48, 75 – 62 | ESLU-Beg | SEC | 16 | L2P, AL |
| | Cecile | 54, 69 – 62 | EXEF | SEC | 23 | L2A, AL |
| | Therese | 57, 68 – 63 | ESLP | ESL-3 | 44 | CL1 Swahili, L2A, PL |
| Advanced | Liliane | 57, 73 – 65 | EXEG | SEC + EE3 | 22 | L2A, AL |
| | Eric | 60, 70 – 65 | ESLU-Adv | SEC | 37 | L2P, GL |
| | Martine | 61, 70 – 66 | UGE | SEC + EE6 | 24 | L2A, GL |
| | Benoit 1 | 60, 74 – 67 | MBA | SEC + EE1 | 23 | L2A, GL |
| | MarieFrance | 63, 72 – 68 | CEGEP | SEC | 17 | L2P, GL |
| | MarieAnnic | 60, 76 – 68 | MILSpouse | SEC + ESL3 | 36 | L2P, GL |
| | Sophie 2 | 60, 76 – 68 | UGE | SEC + IMM | 22 | L2A, GL |
| | Nadine | 67, 68 – 68 | ESLU-Adv | SEC + IMM | 19 | L2A, GL |
| | Melanie | 60, 79 – 70 | UGE | SEC + EE1 | 21 | L2A, GL |
| | Elie | 69, 75 – 72 | UGF | SEC + IMM | 19 | CL1 Arabic, GL |
| | Benoit 2 | 66, 79 – 73 | EXEG | SEC + IMM | 34 | L2A, GL |
| | Greg | 67, 81 – 74 | MBA | SEC + EE1 | 24 | L2A, GL |
| | David | 68, 82 – 75 | PHD | SEC + EE6 | 29 | L2A, GL |
| | Ode | 80, 71 – 76 | UGE | SEC + IMM | 19 | L2A, GL |
| | Justin | 71, 84 – 78 | UGE | SEC + IMM | 19 | L2A, GL |
| | Karlynn | 77, 85 – 81 | UGE | SEC + IMM | 18 | L2P, GL |
| | Erika | 83, 79 – 81 | UGE | SEC + IMM | 19 | L2A, GL |
| | Raphaelle | 77, 85 – 81 | UGE | SEC + IMM | 19 | L2A, GL |

### Abbreviations

***Current professional/educational status***

CEGEP: Enrolled in a Quebec-based, CEGEP
EXEF/G: Executive for a Francophone/Anglophone organization
ESLP: Enrolled in a private ESL school
ESLU-Beg/Adv: Takes beginner or advanced ESL at university
MBA: MBA student on an English campus
MIL(Spouse): Canadian Military (or spouse) - basic Engl. needed
PHD: Full-time PhD student on an English campus
RA: Chemical research assistant (basic English needed)
UGE: Undergrad student on an Anglophone campus

***Other information***

CL1: Concurrent L1 (for participants from Francophone Africa)
GL/AL/PL: Self-rated as good, average or poor L2 listener
L2A/L2P: Self-rated as activ/passive user of L2

***Prior ESL training***

BIT 1-2/14: Back in training after 2 or 14-years
ESL3: ESL training for no more than 3 years
GRAD: Grad studies at Univ. of Alberta
IMM.: Informal immersion in English Canada
SEC: Seconday School Training
TRIPS: Few trips to Anglophone countries
EE 1/3/6: English used everyday for 1/3/6 years

**Appendix G. Word and nonword stimuli for the lexical decision task (Exp. 3)**

*Words*

| Frequency in English | Cognacy with French | Onset Type : C |
|---|---|---|
| HF | YESCOG | case, cash, cause, code, large, list, long, look, made, man, mass, must, news, night, North, note, page, piece, point, post, real, rest, rich, round |
| | NOCOG | coat, could, keep, king, last, late, light, like, mad, mean, mind, moon, need, nice, nine, noise, pick, poor, pull, put, right, road, rule, run |
| LF | YESCOG | cage, calm, cave, cure, lance, lard, lobe, loft, math, mauve, mime, mule, nerve, norm, nude, null, palm, pave, peel, pose, raid, rinse, robe, ruse |
| | NOCOG | cog, couch, keg, kite, kneel, knit, lap, lawn, lick, lid, mend, moan, moose, mop, noon, numb, pad, peep, pip, posh, rash, ripe, rope, rug |

| Frequency in English | Cognacy with French | Onset Type : CC1 |
|---|---|---|
| HF | YESCOG | black, branch, class, crime, glass, group, place, price |
| | NOCOG | blood, brain, close, crowd, glad, great, please, proud |
| LF | YESCOG | blond, brute, clone, crab, gland, grill, plead, prude |
| | NOCOG | blush, bribe, cling, crook, glare, grab, plot, pram |

| Frequency in English | Cognacy with French | Onset Type : CC2 |
|---|---|---|
| HF | YESCOG | skin, smoke, snow, space |
| | NOCOG | school, small, smell, speak |
| LF | YESCOG | skate, smash, sniff, spouse |
| | NOCOG | skull, smear, snooze, spike |

*Nonwords*

Onset Type: C

cath, coaf, coove, coth, kaygue, keab, keff, kife, kish, koun, ladge, laysh, lep, libe, liff, lish, loash, loff, looge, loun, maph, meeve, megg, meitch, mighth, moaf, modge, moofe, moun, myv, nabe, nazz, neam, nemm, nibe, ning, nobe, noome, noude, nozz, pab, payf, peabe, pem, pib, pime, poam, poff, pooth, poun, radge, raithe, reb, rish, riss, roathe, rotch, rouce, ruge, wrygue

Onset Type: CC1

blayk, bleezz, bloud, brett, brike, brode, clayne, cliz, crewk, crose, crouce, gledd, gleed, gloan, grat, grine, grop, klake, plack, pliff, plime, proun, proz, prues

Onset Type: CC2

skal, skaze, skile, skizz, skode, skoon, smag, smake, smit, smot, smout, smoze, sneeff, snell, snett, snite, snouce, snoude, spale, speet, spile, spon, spone, spooze

Onset Type: CC3

fkale, fkett, fpoan, fpous, kfeeze, kfine, lnome, lnool, mzat, mzeev, pnag, pneel, pnice, pnoap, psate, pseet, psine, psoak, psog, psoon, psoosh, psout, ptale, ptetch, ptish, ptoul, srive, srook, tlace, tling, tmoss, tmoul

## Appendix H. Carrier strings for the word-spotting task (Experiment 4)

*Strings carrying HF-YESCOG words*

| Aligned | | Misaligned | |
|---|---|---|---|
| voof.case | [vuf.keɪs] | voo.scase | [vu.skeɪs] |
| dawf.cash | [dɔf.kæʃ] | dou.scash | [daw.skæʃ] |
| jife.cause | [dʒajf.kɔz] | jye.scause | [dʒaj.skɔz] |
| thoaf.code | [θowf.kowd] | thoe.scode | [θow.skowd] |
| chout.large | [tʃawt.lardʒ] | choe.glarge | [tʃow.glardʒ] |
| jate.list | [dʒejt.lɪst] | jay.plist | [dʒej.plɪst] |
| voot.long | [vut.lɔŋ] | choo.plong | [tʃu.plɔŋ] |
| zawt.look | [zɔt.lʊk] | zaw.glook | [zɔ.glʊk] |
| louk.made | [lawk.mejd] | lou.smade | [law.smejd] |
| vooc.man | [vuk.mæn] | foo.sman | [fu.smæn] |
| chike.mass | [tʃajk.mæs] | chye.smass | [tʃaj.smæs] |
| vayk.must | [vejk.mʌst] | vay.smust | [vej.smʌst] |
| fawp.news | [fɔp.njuz] | fye.snews | [faj.snjuz] |
| zayp.north | [zejp.norθ] | zay.snorth | [zej.snorθ] |
| voop.note | [vup.nowt] | thoo.snote | [θu.snowt] |
| thoap.night | [θowp.najt] | voe.snight | [vow.snajt] |
| coaf.page | [kowf.pejdʒ] | koe.spage | [kow.spejdʒ] |
| foof.piece | [fuf.pis] | doo.spiece | [du.spis] |
| douf.point | [dawf.pɔjnt] | daw.spoint | [dɔ.spɔjnt] |
| voaf.post | [vowf.powst] | vou.spost | [vaw.spowst] |
| vaws.real | [vɔs.ril] | vaw.breal | [vɔ.bril] |
| thoas.rest | [θows.rɛst] | thoe.grest | [θow.grɛst] |
| lous.rich | [laws.rɪtʃ] | lou.frich | [law.frɪtʃ] |
| tace.round | [teɪs.rawnd] | tay.bround | [teɪ.brawnd] |

*Appendix H continues next page*

*Strings carrying HF-NOCOG words*

| Aligned | | Misaligned | |
|---|---|---|---|
| fawf.coat | [fɔf.kowt] | faw.scoat | [fɔ.skowt] |
| dife.could | [dajf.kʊd] | thay.scould | [θaj.skʊd] |
| zoaf.keep | [zowf.kip] | fou.skeep | [faw.skip] |
| douf.king | [dawf.kɪŋ] | fou.sking | [faw.skɪŋ] |
| voot.last | [vut.læst] | voo.flast | [vu.flæst] |
| thoot.late | [θut.lejt] | fye.clate | [faj.klejt] |
| choat.light | [tʃowt.lajt] | choe.glight | [tʃow.glajt] |
| vawt.like | [vɔt.lajk] | vaw.clike | [vɔ.klajk] |
| sike.mad | [sajk.mæd] | kye.smad | [kaj.smæd] |
| fooc.mean | [fuk.min] | foo.smean | [fu.smin] |
| vayk.mind | [vejk.majnd] | vay.smind | [vej.smajnd] |
| thoak.moon | [θowk.mun] | thoe.smoon | [θow.smun] |
| vawp.need | [vɔp.nid] | voe.sneed | [vow.snid] |
| zoap.nice | [zowp.najs] | zoe.snice | [zow.snajs] |
| loup.nine | [lawp.najn] | lou.snine | [law.snajn] |
| doup.noise | [dawp.nojz] | dou.snoise | [daw.snojz] |
| fife.pick | [fajf.pɪk] | jye.spick | [dʒaj.spɪk] |
| nawf.poor | [nɔf.pur] | naw.spoor | [nɔ.spur] |
| tafe.pull | [tejf.pʊl] | tay.spull | [tej.spʊl] |
| doof.put | [duf.pʊt] | daw.sput | [dɔ.spʊt] |
| zaws.right | [zɔs.rajt] | zaw.cright | [zɔ.krajt] |
| thoos.road | [θus.rowd] | thoo.croad | [θu.krowd] |
| voas.rule | [vows.rul] | vaw.prule | [vɔ.prul] |
| jous.run | [dʒaws.rʌn] | jou.grun | [dʒaw.grʌn] |

*Appendix H continues next page*

*Strings carrying LF-YESCOG words*

| Aligned | | Misaligned | |
|---|---|---|---|
| zayf.cage | [zejf.kejdʒ] | zay.scage | [zej.skejdʒ] |
| fife.calm | [fajf.kam] | fye.scalm | [faj.skam] |
| poof.cave | [puf.kejv] | poo.scave | [pu.skejv] |
| dawf.cure | [dɔf.kjur] | daw.scure | [dɔ.skjur] |
| fout.lance | [fawt.læns] | fou.blance | [faw.blæns] |
| jate.lard | [dʒejt.lard] | jay.blard | [dʒej.blard] |
| chite.lobe | [tʃajt.lowb] | chye.flobe | [tʃaj.flowb] |
| voot.loft | [vut.lɔft] | chaw.ploft | [tʃɔ.plɔft] |
| sike.math | [sajk.mæθ] | zye.smath | [zaj.smæθ] |
| louk.mauve | [lawk.mov] | lou.smauve | [law.smov] |
| nawk.mime | [nɔk.majm] | naw.smime | [nɔ.smajm] |
| thouk.mule | [θawk.mjul] | thoe.smule | [θow.smjul] |
| zipe.nerve | [zajp.nɘrv] | zye.snerve | [zaj.snɘrv] |
| vape.norm | [vejp.norm] | vay.snorm | [vej.snorm] |
| foop.nude | [fup.nud] | foo.snude | [fu.snud] |
| foap.null | [fowp.nʊl] | thoe.snull | [θow.snʊl] |
| vawf.palm | [vɔf.pam] | vaw.spalm | [vɔ.spam] |
| douf.pave | [dawf.pejv] | dou.spave | [daw.spejv] |
| chife.peel | [tʃajf.pil] | chye.speel | [tʃaj.spil] |
| voof.pose | [vuf.powz] | voo.spose | [vu.spowz] |
| faws.raid | [fɔs.rejd] | faw.fraid | [fɔ.frejd] |
| gous.rinse | [gaws.rɪns] | gou.brinse | [gaw.brɪns] |
| coas.robe | [kows.rowb] | koe.frobe | [kow.frowb] |
| voas.ruse | [vows.ruz] | voe.pruse | [vow.pruz] |

*Appendix H continues next page*

*Strings carrying LF-NOCOG words*

| Aligned | | Misaligned | |
|---|---|---|---|
| jafe.cog | [jejf.kɔg] | vay.scog | [vej.skɔg] |
| voof.couch | [vuf.kawtʃ] | choe.scouch | [tʃow.skawtʃ] |
| zife.keg | [zajf.kɛg] | zye.skeg | [zaj.skɛg] |
| poaf.kite | [powf.kajt] | poe.skite | [pow.skajt] |
| thoop.kneel | [θup.nil] | thoo.sneel | [θu.snil] |
| choap.knit | [tʃowp.nɪt] | choe.snit | [tʃow.snɪt] |
| vawt.lap | [vɔt.læp] | vaw.blap | [vɔ.blæp] |
| poot.lawn | [put.lɔn] | voo.clawn | [vu.klɔn] |
| zoat.lick | [zowt.lɪk] | zoe.plick | [zow.plɪk] |
| dite.lid | [dajt.lɪd] | chye.flid | [tʃaj.flɪd] |
| vaik.mend | [vejk.mɛnd] | vay.smend | [vej.smɛnd] |
| dawk.moan | [dɔk.mown] | daw.smoan | [dɔ.smown] |
| thake.moose | [θejk.mus] | thay.smoose | [θej.smus] |
| douk.mop | [dawk.mɔp] | dou.smop | [daw.smɔp] |
| nawp.noon | [nɔp.nun] | naw.snoon | [nɔ.snun] |
| sipe.numb | [sajp.nʌm] | zye.snumb | [zaj.snʌm] |
| fife.pad | [fajf.pæd] | fye.spad | [faj.spæd] |
| zaif.pip | [zejf.pɪp] | zay.spip | [zej.spɪp] |
| koaf.posh | [kowf.pɔʃ] | choe.sposh | [tʃow.spɔʃ] |
| douf.peep | [dawf.pip] | dou.speep | [daw.spip] |
| poos.rash | [pus.ræʃ] | thoo.grash | [pu.græʃ] |
| lous.ripe | [laws.rajp] | lou.cripe | [law.krajp] |
| faws.rope | [fɔs.rowp] | faw.brope | [fɔ.browp] |
| foos.rug | [fus.rʌg] | foo.prug | [fu.prʌg] |

## Appendix I. Stimuli for the main off-line segmentation task (Experiment 6)

*Onset Type: CC1*

jayblard, vawblap, vawbreal, fawbrope, vooflast, chyeflobe, loufrich, fawfraid, choaglight, zawglook, jougrun, thoogrash, vawklike, vooclawn, thoocroad, loucripe, jayplist, chawploft, vawprule, fooprug

*Onset Type: CC2*

fousking, zyeskeg, vaysmind, dawsmoan, voesneed, zyesnumb, nawspoor, fyespad

*Onset Type: CC3*

fawfcoat, jeifcog, choatlight, pootlawn, thoakmoon, doukmop, doupnoise, nawpnoon, faifpick, faifpad, thoosroad, poosrash