

THE CONTRIBUTION OF PHONOTACTIC AND LEXICAL INFORMATION IN THE
SEGMENTATION OF MULTI-WORD UTTERANCES

DISSERTATION

Presented in Partial Fulfillment of Requirements for the Degree Doctor of Philosophy in
the Graduate School of The Ohio State University

By

Lisa Contos Shoaf, M.A.

The Ohio State University

2002

Dissertation Committee:

Professor Mark Pitt, Advisor

Professor Keith Johnson

Professor Shari Speer

Professor James Todd

Approved by:

Advisor

Psychology Graduate Program

ABSTRACT

Locating word boundaries in continuous speech is a complex task that is completed effortlessly by listeners. Determining what sources of information are used by listeners to achieve successful segmentation is critical for developing models of word recognition. Six experiments were designed to look at the use of phonotactic information in speech segmentation. The first three experiments (Part I) replicated and extended the word-spotting findings of McQueen (1998), demonstrating that phonotactic information is used in the segmentation of words (i.e., CVCs) embedded within nonwords. The next three experiments (Part II) were designed to investigate whether phonotactic effects held up in the more naturalistic context of connected speech, where lexical information is also available to guide word segmentation. Phonotactic effects were present, though small, in comparison to lexical effects. Additionally, phonotactic effects generally were weaker in the context of words than in the context of nonwords, although this finding was not always statistically significant. The data suggest a secondary role of phonotactics when in conjunction with lexical information, and also suggest that maximal lexical activation may even diminish the effects of phonotactics in certain circumstances. Implications of these results for future research and for word recognition models are discussed.

ACKNOWLEDGMENTS

I owe a great debt to many people who helped me in my attainment of a Ph.D. First, to my husband Rick who, more than any other, understood my need to achieve this goal. Without his support and encouragement, I could not have succeeded. Second, to my daughter Sydney, who could always manage to make mommy smile when the going got tough. I am also very appreciative of my parents and Rick's parents, who supported my efforts for so many years.

My achievements could never have been possible without the guidance of my advisor, Mark Pitt. He taught me how to be a scientist and a scholar. I consider myself very fortunate to have such a patient, dedicated, and encouraging teacher.

I wish to thank the members of my dissertation committee: Keith Johnson, Shari Speer, and James Todd, for their cooperation and comments. I would also like to acknowledge Heather Trader, who helped run the experiments for this paper.

VITA

August 28, 1971.....born, Akron, Ohio
June, 1993.....B.S in Psychology, The Ohio State University
June, 1997.....M.A. in Psychology, The Ohio State University
1995-2001.....Graduate Teacher/Researcher, The Ohio State University

PUBLICATIONS

Research Publication

1. Pitt, M.A., & Shoaf, L.C. (2001). The source of a lexical bias in the verbal transformation effect. *Language and Cognitive Processes*.
2. Pitt, M.A., & Shoaf, L.C. (2002). Linking verbal transformations to their causes. *Journal of Experimental Psychology: Human Perception and Performance*.

FIELD OF STUDY

Major Field: Psychology

Minor Field: Cognitive/Experimental Psychology

TABLE OF CONTENTS

	<u>Page</u>
Abstract.....	ii
Acknowledgments.....	iii
Vita.....	iv
List of Tables.....	viii
List of Figures.....	ix
Introduction.....	1
Chapters:	
1. Part I: Phonotactic influences on between-word segmentation.....	10
1.1 Experiment 1--McQueen replication.....	10
1.1.1 Method.....	11
1.1.2 Results.....	13
1.1.3 Discussion.....	14

1.2	Experiment 2--Replication of Experiment 1 with lexical decision.....	17
1.2.1	Method.....	19
1.2.2	Results.....	20
1.2.3	Discussion.....	21
1.3	Experiment 3-- Replication of Experiment 2 with two-word utterances.....	23
1.3.1	Method.....	24
1.3.2	Results.....	24
1.3.3	Discussion.....	25
2.	Part II: The interaction of phonotactic and lexical information in between-word segmentation.....	29
2.1	Experiment 4--Interaction of phonotactic and lexical cues.....	32
2.1.1	Method.....	32

2.1.2	Results.....	34
2.1.3	Discussion.....	35
2.2	Experiment 5--Replication of Experiment 4 without the multiple-choice task.....	38
2.2.1	Method.....	38
2.2.2	Results.....	39
2.2.3	Discussion.....	41
2.3	Experiment 6--Testing the limits of phonotactic effects.....	43
2.3.1	Method.....	44
2.3.2	Results.....	45
2.3.3	Discussion.....	46
3.	General Discussion.....	48
	List of References.....	63

Appendices:

A. Stimuli used in Part I.....	67
B. Stimuli used in Part II (Experiments 4 and 5).....	69
C. Stimuli used in Part II (Experiment 6).....	71

LIST OF TABLES

<u>Table</u>		<u>Page</u>
1	Mean RTs for Experiments 1-3 (Part I) as a function of alignment (sd in parentheses). Mean differences between alignment conditions are also shown.....	59
2	Examples of stimuli used in Experiments 1-6.....	60
3	Mean RTs for Experiments 4-6 (Part II) as a function of lexical status and phonotactic status (sd in parentheses). Mean differences between phonotactic conditions are shown in the margins.....	61

LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
1	Display of phonotactic effect size across experiments as a function of lexical status.....	62

INTRODUCTION

Recognizing where a word begins and ends in written English is a simple task for the average reader because clear spaces mark word boundaries. Finding word boundaries in spoken English, however, is anything but simple, as the language provides its listeners with ambiguous cues, or no cues at all, to mark where words begin and end (Cole & Jakimik, 1980). Despite the unfavorable conditions with which a listener must contend, he or she has no difficulty recognizing words in a typical conversation, where words are produced at an average rate of nearly seven syllables per second (Pollack & Pickett, 1964). How the human mind is able to locate word boundaries (i.e., segment speech) in order to recognize words within a continuous speech stream is a question psycholinguistic researchers have asked for decades.

Research on what has been called ‘the segmentation issue’ has focused on those cues that individuals use to recognize words and to identify their boundaries. Over the years, researchers have taken different approaches to studying the segmentation issue. The earliest approach emphasized the role of specific acoustic/phonetic cues to signal word onsets and offsets (Lehiste, 1960; Nakatani & Dukes, 1977). It was thought that such cues would be present in the speech signal and would clearly mark a word’s onset and offset. While intuitively appealing, this data-driven approach had major limitations.

For instance, speech spectrograms reveal that silence, a seemingly obvious word boundary cue, is often absent at word junctures due to coarticulation or is present word-internally prior to the articulation of stop consonants. Allophonic differences (i.e., differences in the way phonemes are articulated as a result of their context and/or position within a word), such as glottal stops, laryngealized voicing, and aspiration can be used to signal word boundaries (Lehiste, 1960; Nakatani & Dukes, 1977). However, these cues to segmentation are not always reliably present in the signal, nor are they sufficient for identifying boundaries (but see Church, 1987).

The unreliable nature of acoustic/phonetic segmentation cues resulted in a shift in focus from data-driven (i.e., bottom-up) approaches to segmentation to knowledge-driven (i.e., top-down) approaches to segmentation. Models that incorporate lexical information as a way to locate word boundaries are interesting because they propose no specific segmentation strategy, but instead consider segmentation a consequence of successful word recognition. In essence, lexical access leads to segmentation. In some of the earliest work on segmentation in continuous speech, Cole and Jakimik (1980) suggested that speech is processed in a strictly left-to-right manner. As a result, as more of a word is heard, its representation increases in activation until it is recognized as the intended word. Once it is recognized, its offset is known, and the onset of the following word is identified. Another influential model, the original cohort model (Marlsen-Wilson & Welsh, 1978), states that words are recognized at the point in which they become unique from all others in the lexicon. Thus, a word such as 'blanket' becomes unique at the 'e' because no other word has the onset /blanjke/. While this seems a reasonable approach to

locating word boundaries, there are problems with its sole use as a cue to segmentation. First, as Luce (1986) pointed out, many words do not become unique at their offset. For example, 'car' can stand alone, or it can be heard as the onset of the longer word 'carpet.' In addition, it is debated whether affixes (e.g., uncover, covering) are included or are stripped off in determining a word's uniqueness point (Wurm, 1997). Finally, some data suggest that uniqueness points no longer serve as useful cues to word identity in rapid continuous speech (Radeau, Morais, Mousty, and Bertelson, 2000).

In addition to strictly sequential approaches to segmentation, many lexically driven approaches also employ competition among lexical candidates (e.g., TRACE: McClelland & Elman, 1986; Shortlist: Norris, 1994) to recognize words, and thus facilitate segmentation. Competition occurs among lexical candidates that are activated simultaneously in the input. Recognition occurs when one candidate is activated to a certain degree above all others. Segmentation results as a by-product of successful word recognition, which is hypothesized to consist of inhibitory and excitatory processes that work in concert to identify the spoken word. A large number of studies and simulations have demonstrated the results of lexical competition on word recognition and segmentation (Goldinger, Luce, Pisoni, & Marcario, 1992; Norris, McQueen, & Cutler, 1995; Pitt & Shoaf, in press; Shillcock, 1990; Slowiaczek & Hamburger, 1992; Vroomen & deGelder, 1995; Zwitserlood, 1989).

To argue that lexical processes alone are sufficient for successful segmentation would be incomplete, however. Phonological properties of the language, such as metrical stress and phonotactics, have also been shown to facilitate segmentation. Metrical stress,

a phonological property based on the patterning of strong and weak syllables in words, has been looked at extensively as a probable source of segmentation information. In English, approximately 90% of content words begin with a strong syllable, which is defined as a syllable having an unreduced vowel (Cutler & Carter, 1987). This type of statistical regularity prompted Cutler and Norris (1988) to create the metrical segmentation strategy (MSS). According to the MSS, individuals use metrical stress as a heuristic to locate word boundaries. Listeners segment the speech stream at the onset of strong syllables, and a new lexical access attempt is made at the onset of each strong syllable. Language acquisition studies (Jusczyk, Houston, & Newsome, 1999) suggest that infants as early as 7.5 months are sensitive to, and can use, strong-weak stress patterns to segment speech. Juncture misperception studies (Cutler & Butterfield, 1992, Vroomen, vanZon, & deGelder, 1996, Experiment 1) showed that listeners are more likely to insert word boundaries before strong syllables (e.g., ‘analogy’ → ‘an allergy’) and delete boundaries prior to weak syllables (e.g., ‘my gorge is’ → ‘my gorgeous’). Word-spotting studies (Cutler & Norris, 1988; Vroomen et. al., 1996, Experiment 2) demonstrated that listeners have difficulty spotting words embedded within two-syllable nonwords where the second syllable is strong when the embedded word spans across the syllable boundary (e.g., ‘mint’ in /minteif/). This is because /teif/ is a strong syllable and triggers a segmentation at its onset, thus disrupting processing of the embedded word and making it harder to spot. When the second syllable was weak (/mintəf/), ‘mint’ was easier to locate because no segmentation attempt was made prior to the onset of /təf/.

Recently, researchers have investigated another phonological property of a language: phonotactics. Phonotactics refers to the rules that govern how phonemes are combined. Each language has its own set of phonotactic rules. For example, in English, the combination [vr] cannot occur together within a syllable, but is a legal combination in Dutch. Phonotactic rules not only specify which combinations of phonemes can co-occur, but also specify where within the syllable the combination can occur. For instance, in English, the combination [sl] can occur in a syllable's onset, but not in its coda. In theory, categorical phonotactic constraints such as these provide a critical cue to the likely location of a word boundary.

Phonotactics can also provide information on less categorical constraints, such as probabilities between sounds within a language. Words consist of concatenations of sounds, and the correlations between co-occurring sounds within a word are greater than the correlations between co-occurring sounds that spread across a word boundary (Hayes & Clark, 1970; Saffran, Newport, & Aslin, 1996). Therefore, it is plausible that a listener might discover word boundaries based on the probabilities of these co-occurring sounds.

The psychological validity of phonotactics has been demonstrated experimentally in studies of word recognition (Massaro & Cohen, 1983; Onishi, Chambers, & Fisher, 2002; Pitt & McQueen, 1998). Church (1987) argues that phonotactics provide 'rich contextual constraints' that can allow for a more efficient search procedure for matching the input to entries in the lexicon. For these reasons, phonotactics has become the focus of much research in speech perception.

Research has shown that phonotactic information is useful in segmentation. In an artificial language study in which multisyllabic speech sequences were repeated, Cowan (1991) had listeners indicate any and all divisions between syllables that were perceived within a test sequence. He found that segmentation of fixed sequences within the larger test sequence was dependent to a great extent on the frequency of repetition of each sequence. Saffran et. al. (1996) briefly exposed listeners (in three 7-minute sessions) to an artificial language made up of six trisyllabic words (created by concatenating twelve English-like CV syllables). The synthetically produced speech stream contained random configurations of these 'words' with no pauses between the items. They found that listeners were able to learn the 'words' contained in the speech stream using only the transitional probabilities (i.e., the probability of co-occurrence) between syllables. In one of the few tests of the on-line use of phonotactic knowledge in segmentation, McQueen (1998) showed that a word embedded in a two-syllable nonsense word was more easily spotted if the embedded word was aligned with a phonotactically imposed syllable boundary, such as [rok] in [fim-rok] (where [mr] is phonotactically illegal in Dutch), than if the embedded word was misaligned with a syllable boundary [fi-drok]. Research on what has been called the possible word constraint (PWC; Norris, McQueen, Cutler, & Butterfield, 1997) suggests that when listeners attempt to segment a speech stream, they are not likely to try segmentations that leave phonotactically illegal combinations stranded, such as the 'f' in 'fapple'. van der Lugt (2001), using both word-spotting and lexical decision experiments, looked at whether distributional cues are used in speech segmentation. He pitted high- and low-

probability CV or VC pairs and found that a high-probability sequence at the onset of a word made the word easier to spot. Taken together, these data provide solid evidence that phonotactic information is used in segmentation.

Phonotactics: Another look

Intuitively, the idea of phonotactic information as a cue to speech segmentation is intriguing. Because phonotactic constraints govern how phonemes combine, it is an inherent property of naturally-occurring speech. Experimentally, the few studies previously mentioned that have manipulated phonotactics suggest that it is a useful segmentation cue.

Closer inspection of these studies suggests that our knowledge regarding phonotactic cues to segmentation is limited to within-word segmentation. The word-spotting studies discussed earlier investigated the effects of phonotactics in locating a word contained within a longer nonword (e.g., 'rok' in 'fimrok'). In essence, these were investigations of the influence of phonotactics on identifying syllable boundaries within isolated words, not investigations of the influence of phonotactics on identifying boundaries between words. For successful communication, listeners must be able to segment between words, not within words. This difference is a subtle, yet important one. Syllable and word boundaries are highly correlated; however, research suggests that there are acoustic/phonetic cues that can signal word boundaries which may not be present at syllable boundaries (Church, 1987; Davis, Marslen-Wilson & Gaskell, 2002; Gow & Gordon, 1995; Lehiste, 1960). Thus far, there is little experimental evidence to suggest

that the usefulness of phonotactics can be extended to segmentation between words. While the artificial language studies (Cowan, 1991; Saffran et. al., 1996) take a step in this direction by investigating listeners' ability to locate boundaries between 'words', the stimuli in the study were created so as to be devoid of naturally occurring and potentially useful segmentation cues (e.g., acoustic/phonetic, lexical). Can the prior findings be generalized to the segmentation of words in an environment similar to what listeners normally encounter, where other cues are present to assist in between-word segmentation?

As discussed earlier, lexical information has been shown to be a powerful cue to achieving successful segmentation. In fact, many lexically-based models of word recognition claim to achieve successful segmentation based on lexical processes (i.e., excitation and inhibition of lexical items) alone. If lexical cues to segmentation, in addition to acoustic cues, are already present in connected speech, is phonotactic information even necessary? How influential could phonotactic information be? Will both sources equally influence word segmentation? Will one source be more influential than the other source? The few studies that have investigated multiple-cue interaction in speech segmentation (Christiansen, Allen, & Seidenberg, 1998; Johnson & Jusczyk, 2001; Norris et. al., 1995; Vroomen, Tuomainen, & de Gelder, 1998) suggest that cues can interact in complex ways. Exactly how phonotactic information combines with lexical information in segmenting connected speech is the focus of the experiments described in this paper.

The goal of the present study is to bring to light the nature of the interplay between phonotactic information and lexical information in between-word segmentation. The starting point of this study is McQueen (1998). McQueen's word-spotting experiments demonstrated quite convincingly that phonotactic information is used by the listener in locating syllable boundaries. When the embedded word was in line with a phonotactically imposed syllable boundary (e.g., [rok] in [fim-rok]), it was easier to spot than when the embedded word was misaligned by one segment with a phonotactically imposed syllable boundary (e.g., [rok] in [fi-drok]). These results were found regardless of whether the embedded word occurred at the beginning of the nonsense sequence or at the end of the nonsense sequence, although the effects were larger when the embedded word was located at the end of the sequence. The question remains whether these findings can be extended to show listeners' use of phonotactic information to locate boundaries between two words in a context that approximates connected speech.

The first three experiments (Part I) generalize from a replication of McQueen's (1998) word-spotting experiment to a within-context lexical decision task, designed to study between-word segmentation. The aim of these experiments is to demonstrate the reliability and the sensitivity of the new task in uncovering the influence of phonotactic information in segmenting words. Part II assesses whether the phonotactic effect still holds up in segmentation between words, where lexical information is also present.

CHAPTER 1

PART I: PHONOTACTIC INFLUENCES ON BETWEEN-WORD SEGMENTATION

1.1 Experiment 1--McQueen (1998) replication

The goal of Experiment 1 was to replicate the results of McQueen (1998) in English. McQueen's study was conducted in Dutch, which has different phonotactic restrictions than does English, one of which is that certain phonemes cannot occur in a syllable-final position (e.g., /d/ cannot occur in the coda position of a syllable). Because of these cross-language differences, it was important to show that the same effects could also be found in English. Modifications were made to the design, in part to accommodate the differences between the languages. The structure of the embedded words in the McQueen experiment were all CVC. The structure of the embedded words in the current experiment was less restricted: they were either CVC, CVCC or CCVC. As in the McQueen study, in the aligned condition, the embedded word was aligned with a phonotactically mandated boundary of a nonsense sequence (e.g., 'case' in 'sibcase', where [b-k] requires a boundary). In the nonaligned condition, there was no phonotactically mandated boundary with which to align the embedded word (e.g., 'case' in 'siscase,' where [s-k] does not require a phonotactic boundary between syllables). In the comparable condition in the McQueen study, the embedded word was misaligned by

one segment with a phonotactically mandated boundary (e.g., [rok] was embedded in [fi-drok], where [d] never occurs in the syllable-final position). The slight modification made to the stimuli in the present experiment (because of cross-language differences) should not impact the results if phonotactic information is a useful segmentation cue.

The predictions are straightforward. If phonotactic information guides segmentation, then when the embedded word is aligned with a phonotactically imposed syllable boundary (aligned condition: e.g., ‘case’ in ‘sibcase’), it should be easier to spot than when the same embedded word is not aligned with a phonotactically imposed syllable boundary (nonaligned condition: e.g., ‘case’ in ‘siscast’). This will be demonstrated in faster response times (RTs) in the aligned condition than the nonaligned condition.

1.1.1 Method

Participants. Twenty-four introductory psychology students at The Ohio State University participated in the experiment for class credit. All were native English speakers born in the US and reported normal hearing.

Stimuli. Forty-eight pairs of target stimuli were created. Each pair consisted of a target word (median frequency=58.5) embedded word-finally in one of two different types of nonword carriers. In the aligned condition, phonotactic constraints required a syllable boundary be placed before onset of the embedded word (e.g., ‘sib-case’). In the nonaligned condition, the phonotactic constraints did not require a syllable boundary be placed prior to the onset of the embedded word (e.g., ‘sis-case’). Table 1 shows a sample

of the stimuli used in this experiment as well as the other experiments. All stimuli were produced with a full vowel in each syllable. Twenty-four pairs had embedded words which began with a liquid (/l/ or /r/), as did all of McQueen's stimuli. The other 24 pairs had embedded words which began with a stop (/k/, /p/, /t/) or a nasal (/m/ or /n/). These stimuli were included to show that McQueen's results could be replicated and extended to a wider variety of stimuli. Another 48 nonwords were created as filler items. The filler items had the same general structure as the target items but contained no embedded words. See Appendix A for a list of all target stimuli.

The targets and the fillers were used to create two lists. Each list contained all 48 target items (half in the aligned condition, half in the nonaligned condition) and all 48 filler items. Twenty-four of the target items in each list had stop/nasal-onset embedded words, and the other 24 target items had liquid-onset embedded words. The list was pseudo-randomized, such that there were no long strings of target items. In addition, five 'warmup' filler trials were created, three of which had embedded words. These were presented as the first five trials of the test trials. A set of twenty practice trials were also made in which half of the trials contained embedded words.

Procedure. All materials were recorded by the author onto DAT tape, sampled at 48 kHz and down-sampled to 16 kHz. The stimuli were then transferred to the hard drive of a PC, where they were later edited and saved as individual sound files (.wav format). To keep the embedded word constant across the aligned and nonaligned conditions, splicing was performed. The embedded word was spliced out of the nonsense word in the aligned and the nonaligned condition, and a single token of the embedded word

(whichever of the spliced tokens sounded better) was then pasted back into the nonwords of each of the two conditions, thereby creating an aligned and nonaligned nonword containing the identical embedded word. PC hardware/software controlled the stimulus presentation, timing, and data collection.

Participants were seated, up to four at a time, in individual sound attenuated booths. They were told that on a given trial they would hear a single nonsense word which may or may not have a hidden word located within it. As with McQueen (1998), they were told that the hidden word may occur toward the end of the nonsense word. Their task was to listen for an embedded word, and to press a button labeled 'word' only if a real word was heard. After pressing the 'word' button, they were instructed to write down the word that they spotted on an answer sheet which consisted of columns of blank lines numbered from 1-101 (five warmup and 96 test trials). If they did not spot a real word within the nonsense word, they were told to simply put a line through that trial number on the answer sheet. To help them keep their place, the trial number appeared on a computer monitor for the entire duration of that trial. Participants had a total of 10 seconds to press the button and to write down the hidden word. The entire experiment lasted about 20 minutes.

1.1.2 Results

The response times (RTs) of 24 participants were collected in the aligned and nonaligned conditions for both stimulus types and the means are displayed in the top third of Table 2. RTs were calculated in the same manner as McQueen (1998), by subtracting

the duration of the entire stimulus from the raw RT (measured from stimulus onset). This resulted in a measurement of RT from word offset. Overall, RTs were 82 ms faster in the aligned condition than the nonaligned condition. This difference was reliable by subjects, $F(1,23)=9.20$, $p=.006$, and by items, $F(1,46)=4.96$, $p=.031$.

Next, error data were analyzed. Of the 21 participants who made errors (misses only), errors were over twice as great in the nonaligned condition (mean=3.81) than in the aligned condition (mean=1.57). This difference was significant by subjects, $F(1,20)=14.84$, $p=.001$, and by items, $F(1,47)=12.67$, $p=.001$.

Finally, the data were analyzed as a function of stimulus type (i.e., stops/nasals vs. liquids). As a whole, stop/nasal onset embedded words were responded to significantly (105 ms) faster than liquid-onset embedded words, a difference which was reliable by subjects, $F(1,23)=12.54$, $p=.002$, as well as by items, $F(1,46)=7.84$, $p=.007$. Despite the large differences in RT, neither the ANOVA by subjects nor by items showed an interaction, $F<1$. Rather, both sets of stimuli showed the same trend: RTs to aligned embedded words were faster than RTs to nonaligned embedded words. For the liquid-onset embedded words, the aligned condition was 104 ms faster than the nonaligned stimuli. For the stop/nasal-onset embedded words, the aligned condition was 60 ms faster than the nonaligned condition.

1.1.3 Discussion

Experiment 1 successfully replicated the word-spotting results of McQueen (1998) for words embedded at the end of nonsense sequences. The results clearly show

that listeners use phonotactic information in detecting embedded words. Strong alignment effects were shown both in RTs and errors. When an embedded word was aligned with a phonotactically imposed syllable boundary (as in 'sib-case'), it was easier to detect than when the embedded word was not aligned with a phonotactically imposed syllable boundary (as in 'sis-case'). This suggests that listeners are aware of and make use of the phonotactic constraints of English.

One interesting and unexpected finding was the difference between word-spotting RTs for liquid-onset and stop/nasal-onset embedded words. Overall word-spotting RTs were significantly faster for stop/nasal-onset embedded words than for liquid-onset embedded words. In addition to the global RT difference between stimulus types, the phonotactic effect was smaller in the stop/nasal-onset embedded condition than in the liquid-onset embedded condition.

What accounts for the differences between the two types of stimuli? The mean difference (104 ms) between the aligned and nonaligned conditions found in the liquid-onset stimuli closely matches that found by McQueen (94 ms), who also used liquid-onset stimuli. The smaller difference (60 ms) between the aligned and nonaligned conditions in the stop/nasal onset stimuli, while still significant, suggests that listeners may have had other sources of information available to them to assist in locating word onsets.¹

Analysis of the spectrograms of the stop-onset and liquid-onset stimuli showed that the stop-onset stimuli contained a potentially meaningful period of silence between the offset of the first syllable and the onset of the embedded word (e.g., bis-cast). The mean duration of this silent gap was 59 ms. This silent gap was not present in the liquid-

onset stimuli, nor was it present in the nasal-onset stimuli. If the presence of this silent gap (or the presence of any other acoustic/phonetic cue for that matter) makes stops more salient or easier to process, perhaps phonotactic information may not be as critical to word-spotting as it is in the liquid-onset condition, resulting in a smaller, yet still significant RT difference between the aligned and nonaligned conditions. In addition, this silent gap may contain place of articulation information, allowing listeners to begin processing the stop consonant before it is actually articulated (Keith Johnson, personal communication, July 7, 2002). This may also explain the faster RTs reported for the stop-onset stimuli.

A second possibility has to do with the splicing technique used to create the aligned and nonaligned stimuli. Perhaps splicing introduced artifacts into the liquid-onset stimuli, thus making word-spotting more difficult in the liquid-onset embedded words. That the stimuli were spliced does not compromise the fact that listeners were able to make use of phonotactic information in locating the onset of the embedded words.

Despite the differences obtained across stimulus types, the goal of Experiment 1 was achieved. The experiment successfully replicated the results of McQueen (1998) and extended his findings to English and to a wider variety of stimuli.

1.2 Experiment 2--replication of Experiment 1 with lexical decision

Experiment 1 was critical for demonstrating that the word-spotting results found by McQueen (1998) could be extended to English and to a broader range of stimuli. The next step is to investigate whether the phonotactic effects found in Experiment 1 would hold in a study of between-word segmentation. The word-spotting task had to be modified, as it was designed to investigate issues of within-word segmentation. Word-spotting requires the listener to segment words out of nonsense contexts (i.e., the carrier nonword within which they are contained), and therefore was not appropriate for studying segmentation between words in a real-word context. Before using the new experimental setup to study between-word segmentation, it was necessary to first demonstrate that this new task was sensitive to phonotactic effects. In order to be assured that the new setup was sensitive to phonotactic effects, the same stimuli that were used in Experiment 1 were used in this experiment. If the new experimental setup is sensitive to phonotactic effects, then the results of Experiment 1 should be found in this experiment.

In the new experimental setup, participants listened to multi-word utterances and made lexical decisions to the second item in the utterance, the target word (TW). The speed of their response to the TW was taken to be an indication of the ease with which they were able to successfully segment the item from its context. Listeners who correctly and quickly made a lexical decision to the TW were assumed to have easily located the word boundary between the first word (called the critical word, or CW) and the TW, and

thus were able to easily identify the onset of the TW. Listeners who slowly or incorrectly (or never) made a lexical decision to the TW were assumed to have had difficulty in finding the word boundary between the CW and the TW.

To assess the role of phonotactic information in segmentation, the CW was manipulated in ways intended to influence how easily it could be identified. The boundary consisting of the offset of the CW and the onset of the TW was either a legal combination (such as [r-k], ‘car-case’) or an illegal combination (such as [n-k], ‘van-case’). The results of Experiment 1 suggest that a phonotactically imposed boundary, as in an illegal phoneme combination, will make segmentation easier, resulting in a fast lexical decision RT to the TW. If a boundary is not phonotactically required, as in a legal phoneme combination, listeners will have a harder time segmenting the items, resulting in a slow lexical decision RT to the TW.

This experimental setup has some potential benefits over other tasks previously used to study segmentation. First, and most obvious, the task is being performed on connected speech, which is more representative of natural speech than isolated words. Second, this is an on-line task (i.e., a task in which speeded responses are a requirement), which means that RTs are likely to reflect raw processing times, free of metalinguistic knowledge. Third, by manipulating the CW and keeping the TW constant across conditions, the task allows for comparisons to be made between RTs to the same TW across the different conditions.

The purpose of Experiment 2 was to demonstrate that this new setup can measure phonotactic effects. In order to be assured that any effects found are due to a change in

task, and not a change in stimuli, it was important to use the same stimuli that were used in Experiment 1. Recall that these stimuli were produced as a single utterance (e.g., ‘sibcase’). Similar trends should emerge if the task is sensitive to phonotactic constraints.

1.2.1 Method

Participants. Thirty-six participants meeting the same criteria as those in Experiment 1 were used.

Stimuli. The same tokens of the 48 pairs of two-syllable target nonwords that were used in Experiment 1 were also used as target items in this experiment. To bias participants to perceive the stimuli as consisting of two items (e.g., ‘sib case’) rather than as a single item as they were produced (‘sibcase’), two-word filler trials of similar structure to the target words (e.g., CVC CVC or CVC CCVC) were created. The filler items consisted of 12 word-word pairs, 48 word-nonword pairs, and 12 nonword-nonword pairs. The fillers were recorded in the same manner as the target pairs were in the previous experiment, except that the fillers were produced as two words rather than as two-syllable words. Care was taken to make the filler items closely match the target items in speaking rate and overall intonation.

Each of the two lists contained 24 aligned targets, 24 nonaligned targets, and 72 filler items, in a completely randomized order. In addition, 16 practice trials were created which consisted of the same word-nonword combinations as the fillers.

Procedure. Participants were tested in individual sound-attenuated booths. They were told that they would always hear phrases consisting of two items, and that these two items could be any combination of words and/or nonwords. They were instructed to listen carefully to the entire phrase and to make a lexical decision to the second item of the phrase. Fast responding was emphasized. No mention was made that some of the stimuli were actually produced as single two-syllable words rather than two single-syllable words. Participants were given up to four seconds to respond before the next trial began. The entire experiment lasted less than 20 minutes.

1.2.2 Results

The responses of 34 participants were used in the analyses. Two participants' data were excluded for high (>15%) overall error rates.

RTs were measured a bit differently than they were in Experiment 1. This time RTs were measured from the onset of the embedded word (because the embedded words were of equal durations across alignment conditions within each stimulus set) rather than from the offset of the embedded word.

RTs were collected in the aligned and nonaligned conditions and are displayed in the middle rows of Table 2. Overall, RTs were 78 ms faster in the aligned condition than in the nonaligned condition. This difference was significant by subjects, $F(1,33)=11.89$, $p=.002$, and by items, $F(1,47)=7.48$, $p=.009$.

Next, error data were analyzed. As with Experiment 1, there was a large difference between the aligned and nonaligned conditions. Of the 27 participants who

made errors, the errors occurred nearly four times as often in the nonaligned condition (mean=5.22) than in the aligned condition (mean=1.41), which was significant, $F(1,26)=104.17$, $p=.00$. Collapsed over participants, an analysis of the 35 stimuli in which errors were made also showed that errors were much more frequent in the nonaligned condition (mean=4.20) than in the aligned condition (mean=1.09). This was significant, $F(1,34)=21.88$, $p=.00$.

When analyzed as a function of stimulus type, the data patterned similarly to that which was found in Experiment 1. Stop/nasal onsets produced phonotactic effects which were smaller (yet still significant) than liquid-onsets. The difference between stimulus types was significant both by subjects, $F(1, 33)=13.96$, $p=.001$ and by items, where stimulus type was a between-items factor, $F(1,46)=6.50$, $p=.014$.

1.2.3 Discussion

The goal of Experiment 2 was to replicate the findings of Experiment 1 using a new task. This was accomplished. Overall, there was an effect of alignment, such that words aligned with a phonotactic boundary were responded to faster than words that were not aligned with a phonotactic boundary. This was found for both the RT data and the error data. In addition, the size of the phonotactic effect was comparable across experiments. The experiment thus demonstrated that the lexical decision task can and does yield the same results as the word-spotting task.

Similar trends emerged when the data were analyzed as a function of stimulus type. As in Experiment 1, the effect size was larger for liquid-onset stimuli. Also, like

Experiment 1, RTs were slower overall for the liquid-onset stimuli. This replication is not surprising given that the same stimuli that were used in Experiment 1 were also used in this experiment.

1.3 Experiment 3--replication of Experiment 2 using two-word utterances

Experiments 1 and 2 demonstrated that phonotactic information is used in locating boundaries within words. The next step is to test whether phonotactics can be used to locate boundaries between words, where cues to word boundaries, not just syllable boundaries, are present. The stimuli of Experiment 1 and 2 were spoken as single utterances; therefore, they lacked any acoustic/phonetic information that may normally be present in the signal to indicate word boundaries. The goal of Experiment 3 was to test whether the findings of Experiment 2 replicate when word-boundary cues are present as opposed to only syllable-boundary cues.

To investigate the usefulness of phonotactics in locating word boundaries, the target stimuli were re-recorded as two single-syllable utterances (e.g., 'sibcase' was re-recorded as 'sib case'). The target word (TW, formerly the embedded word) was spliced into the aligned and nonaligned conditions in the same manner as described in Experiment 1 (e.g., 'case' was spliced into 'sis case' and 'sib case') to minimize the chance that acoustics differentially influence the aligned and nonaligned conditions. Experiment 2 was then rerun using the new stimuli.

The predictions are the same as those for Experiment 2. If listeners are sensitive to phonotactics, then the presence of an illegal phoneme combination at the word boundary (aligned condition, e.g., 'sib case') should make it easier for listeners to segment the utterance, resulting in a fast lexical decision RT to the second word ('case'). In contrast, the presence of a legal phoneme combination at the word boundary (nonaligned condition, e.g., 'sis case') should make segmentation more difficult, resulting

in a slow lexical decision RT to the second word. Replication of the results of Experiment 2 would suggest that phonotactic information can be used in addition to any acoustic/phonetic information present in the utterance as a result of it being produced as two words. If the results do not replicate, this would suggest that the presence of acoustic/phonetic cues to segmentation can diminish the influence of phonotactics.

1.3.1 Method

Participants. Forty students meeting the same criteria as the previous experiments participated.

Stimuli. Different tokens of the same 48 pairs of target items used in Experiments 1 and 2 were used in this experiment. They were re-recorded in a continuous fashion as two single-syllable words. An attempt was made to keep the speaking rate, amplitude, and intonation the same throughout recording of each stimulus set to allow for easier splicing of the target word within each target pair. Recording, editing, and splicing were performed in the same manner as described in Experiment 1. The filler trials and practice trials that were used in Experiment 2 were also used in this experiment. Lists were also identical to those used in Experiment 2.

Procedure. The procedure was identical to Experiment 2.

1.3.2 Results

The responses of thirty-eight participants were analyzed. Two participants' data were discarded because their overall error rate exceeded 15%. As with Experiment 2, RTs were measured from the onset of the TW.

Mean RTs in the aligned and nonaligned conditions are displayed in bottom third of Table 2. Some striking differences immediately come to light. Overall, an ANOVA by subjects and by items showed essentially no difference between the aligned and nonaligned conditions, $F(1, 37) < 1$, $F(1, 46) = 1.03$, $p = .316$, respectively. Likewise, the subject and item ANOVA revealed no differences between the stops/nasal-onsets and the liquid-onsets, $F < 1$.

One must be careful in interpreting these findings in light of the large interaction of alignment and stimulus type, $F(1, 37) = 8.84$, $p = .005$, $F(1, 46) = 8.20$, $p = .006$. Liquid-onsets and stop/nasal onsets essentially showed the opposite effect. Stop/nasal onsets showed a phonotactic effect similar to what had been found in the previous two experiments: The aligned condition was 52 ms faster than the nonaligned condition. Liquid onsets showed the reverse effect: the aligned condition was 50 ms *slower* than the nonaligned condition. The two results essentially cancel each other out, resulting in no main effect for either alignment or type. Errors to target items were very infrequent, ranging from 1.8% in the aligned condition to 2.5% in the nonaligned condition.

1.3.3 Discussion

The goal of Experiment 3 was to demonstrate that phonotactic information can aid between-word segmentation when acoustic/phonetic cues to word boundaries are present in the signal. The results showed that this was the case for the stop/nasal-onsets. For these stimuli, the aligned condition was significantly faster than the nonaligned condition, an effect that was also found in Experiment 2. This phonotactic effect was found even in the presence of an acoustic word-boundary cue: silence. Recall that an acoustic analysis of the stop-onset stimuli in the two-syllable utterance (e.g., 'sibcase', Experiment 1 and 2) showed the presence of a silent gap between the two syllables, a possible syllable-boundary cue. This cue was not present in the nasal-onset or in the liquid-onset stimuli. Acoustic analysis of the stop-onset stimuli produced as two-word utterances ('sib case', Experiment 3) showed that the average duration of the silent gap between words increased from 59 ms to 82 ms. Converted to a proportion of the overall duration of the stimulus (to control for speaking rate differences across the two recording sessions), the between-syllable gap proportion was .067, whereas the between-word gap proportion was .092. This increase was significant, $F(1,37)=38.07$. Despite the presence of a potential acoustic segmentation cue, phonotactic information was still shown to be useful in between-word segmentation.

The data from the liquid-onset stimuli are more complex. Compared to Experiment 2, average RTs sped up a fairly consistent amount across three of the four conditions: 61 ms for the stop/nasal-aligned stimuli (e.g., 'sib case'), 45 ms for the stop/nasal-nonaligned stimuli (e.g., 'sis case'), 42 ms for liquid-aligned conditions (e.g.,

‘chiyn rake’), and a whopping 212 ms in the liquid-nonaligned condition (e.g., ‘chiyk rake’). In general, the speedup suggests that the addition of acoustic/phonetic word boundary information was helpful to listeners. Why the speedup was so drastic in the liquid-nonaligned condition is unclear, however. This speedup resulted in a reversal of the phonotactic effect that had been consistently found in the previous experiments.

Given that the only difference between Experiment 2 and Experiment 3 was the addition of word boundary cues, further acoustic analyses were performed comparing the two-syllable stimuli with the two-word stimuli, with the hopes of shedding some light on the reversed phonotactic effect. An obvious acoustic cue that was investigated was the prosodic contour of the stimuli. Perhaps when the stimuli were produced as two words rather than as two syllables, a prosodic boundary tone was used to cue the presence of a new word. An analysis of the fundamental frequency contour² (an indication of the use of a prosodic cue to signal a word boundary) does show a difference between overall F0 contours when the stimuli were produced as two words (Experiment 3) and when they were produced as two syllables of a single word. This was found regardless of stimulus type. This difference was significant, $F(1,47)=15.47$, $p=.00$. This potentially useful word-boundary cue could explain the overall RT speedup from Experiment 2 to Experiment 3.

Why would the phonotactic effect change direction in the liquid-onset case though? Perhaps the phonemes /l/ and /r/ become more salient when they are produced in a word-onset context. Nakatani and Dukes (1977) found that the word-initial allophones of /l/ and /r/ are strong cues for word boundaries. If specific acoustic/phonetic properties

of the liquid-onset stimuli are sufficient for locating word boundaries, then maybe phonotactic information is not necessary for these items. If the word-initial /l/ and /r/ allophones strongly cue a word boundary, then the phonotactic effect may simply disappear or even switch directions if, for instance, additional acoustic/phonetic information in the nonaligned condition is present to specify the boundary location. It is difficult to imagine, though, that the acoustic/phonetic features of the liquids are more influential in locating between-word boundaries than the large silent gaps that precede the stop consonants in the stop-onset stimuli (which still showed a phonotactic effect).

What do these results say about McQueen's (1998) data? When the investigation of phonotactic effects was broadened to include between-word segmentation, the phonotactic effect disappeared, a result not anticipated by his word-spotting data. This suggests that phonotactic effects may be limited to instances in which acoustic/phonetic cues are insufficient, as in the case of within-word (i.e., syllable) segmentation.

Given the questionable nature of these stimuli in the between-word context, it was decided to discontinue the use of the liquid-onsets in the future experiments. Because the stop-onset stimuli reliably showed phonotactic effects in within-word and between-word segmentation, these stimuli were used in the experiments of Part II.

CHAPTER 2

PART II: THE CONTRIBUTION OF PHONOTACTIC AND LEXICAL INFORMATION IN BETWEEN-WORD SEGMENTATION

The experiments of Part II move one step closer to a continuous speech context. The goal of the experiments is to determine whether phonotactic effects hold up in the context of multi-word utterances, where lexical information is also available to guide segmentation. The experimental setup remained the same as that used in Experiment 3: participants were exposed to a multi-word utterance consisting of a critical word (CW) followed by a target word (TW), and made a lexical decision to the TW. Some important changes were made to the stimuli, however. First, the TWs consisted of only stop, nasal, and fricative onsets. The results of the experiments in Part I demonstrated that these stimuli offered the best chance of finding reliable phonotactic effects. Second, in addition to a phonotactic manipulation, a lexical manipulation was also included. The CW was either an English word or a nonword. Third, the length of the multi-word utterances was increased from two to three items in length. The additional item always occurred prior to the CW, so that the CW was completely bounded by speech, rather than the being bounded on one side by silence (an obvious segmentation cue!). Fourth, the terminology was changed slightly. Instead of calling the target word (TW) aligned or

nonaligned, the phonotactic conditions were either legal continuations or illegal continuations. In other words, the offset of the CW could either continue into the onset of the TW (e.g., ‘car case’) or the offset of the CW could not, due to phonotactic constraints, continue into the TW (e.g., ‘tan case’).

Crossing the phonotactic and lexical factors created four CW conditions: legal-word, illegal-word, legal-nonword, illegal-nonword. The legal- and illegal-nonword conditions (e.g., ‘nice var case’, ‘nice zan case’) were essentially replications of Experiment 3 (e.g., ‘sib case’, ‘sis case’) with an additional word preceding the CW. I predict that the results of these conditions will pattern similarly to those found in Experiment 3, with RTs in the illegal-nonword condition being faster than RTs in the legal-nonword condition. The addition of lexical information in the legal- and illegal-word conditions (e.g., ‘nice car case’, ‘nice tan case’) could result in several interesting outcomes. A main effect of lexicality would not be surprising, given the large amount of research demonstrating its usefulness in achieving segmentation. A main effect of phonotactics is also possible, as the experiments in Part I have consistently demonstrated a phonotactic effect. Combining these two cues could result in the fastest RTs to the illegal-word condition (which has both the phonotactic and the lexical advantage), and the slowest RTs to the legal-nonword (which has neither advantage).

Whether lexical and phonotactic cues interact, however, is uncertain. One could imagine an interaction between the two cues, such that the presence of one cue minimizes the effect normally found with the other cue. Given the findings of Experiment 1 and 2, where acoustic/phonetic information appeared to mediate the phonotactic effect for stop-

onsets, and Experiment 3, where acoustic/phonetic cues may have mediated the phonotactic effect for liquid-onsets, the interaction of lexical and phonotactic information is not an unreasonable prediction.

2.1 Experiment 4--phonotactic and lexical manipulation

If phonotactic information aids listeners in segmenting words, will the effect still hold in context, where lexical information is also available to influence segmentation? Experiment 4 was designed to test this idea. It is predicted that the presence of both phonological and lexical information will guide segmentation, resulting in fast RTs to the TW. Exactly how these cues will combine to influence segmentation is unclear. The results of this experiment will help provide an answer to this question.

2.1.1 Method

Participants. Forty-two students from the same pool and with the same criteria as the previous experiments participated.

Stimuli. Forty-eight sets of four stimuli were chosen. Each utterance within the set consisted of three items, Word0 (W0, a filler word), followed by the critical word (CW) and the target word (TW). The CW was the item which was manipulated lexically and phonotactically (at its offset). The TW was the item to which listeners made a lexical decision. For target trials, this item was always a single syllable noun or verb. W0 was always a high-frequency single-syllable word. The four conditions that resulted from the crossing of the two factors: legal-word, illegal-word, legal-nonword, and illegal-nonword, were always preceded by the same single-syllable W0 (e.g., ‘nice car case’, ‘nice tan case’, ‘nice var case’, ‘nice zan case’). In all but two cases, the illegal boundary phonemes (i.e., the phoneme at the offset of the CW+the phoneme at the onset of the

TW) were illegal both as word-final combinations and as word-initial combinations. In other words, combinations such as /nk/ (not /ŋk/), which cannot occur as a word-final combination (i.e., /tænk/ is illegal), nor as a word-initial combination (i.e., /nkes/ is illegal) were used. The two exceptions (both /sm/) were combinations that are illegal word-finally but legal word-initially. See Appendix B for a complete list of target stimuli.

The phrases within each phonotactic set were created to maximize phonemic similarity in the boundary region between the two words (e.g., ‘tan case’, ‘zan case’; ‘car case’ ‘var case’) in order to minimize the effect of other acoustic/phonetic cues influencing their segmentation. In addition, the phrases were created to be as semantically neutral as possible, given the constraints of the design.

In addition to using the 48 target sets of stimuli, 152 filler items were also created. The filler stimuli were varied along four dimensions in order to reduce stimulus predictability: phrase length (3-5 words), lexical status of the CW, lexical status of the TW (because the task is lexical decision), and syllabic length of the CW.

All stimuli were recorded onto digital audio tape at a 48 kHz sampling rate (downsampled to 16 kHz) and were digitally transferred to the hard disk of a PC, where they were edited and saved as individual sound files. LCS recorded all four phrases for each target set before moving on to the next set of stimuli, to ensure that the phrases within each target set were similar in terms of speaking rate and loudness. In addition, each phrase was spoken within the larger sentence, ‘The phrase is next,’ to ensure that

the overall intonation pattern remained fairly constant across all stimuli. The phrases were later excised from the longer sentences.

The stimuli were divided into four lists. All 48 TW appeared in each list (e.g., ‘case’ was found in each list), while the CW varied across lists (e.g., ‘nice tan case’, ‘nice zan case’, ‘nice car case’, and ‘nice var case’ occur in separate lists). 152 fillers were also included in each list, for a total of 200 trials. Twenty practice trials of similar structure to those found in the experiment were also included.

Procedure. Up to four participants at a time were tested in individual sound-attenuated booths. They were instructed that they would hear a phrase consisting of two, three, or four items. These items could be either words or nonwords. They were told to listen to the entire phrase and to identify the third word of the phrase as either a real English word or a nonsense word. Fast responding was stressed. Examples of phrases of varying length and of varying word-nonword combinations were then given to ensure that the participants understood the nature of the stimuli and of the task.

The participants first engaged in a twenty-item practice session, where their responses were monitored by the experimenter. After the practice trials were completed, the experiment began. The first twenty experimental trials were ‘warmup’ trials which consisted of all filler trials. Participants had four seconds in which to make a response. Because of the short duration of the experiment, no breaks were given.

2.1.2 Results

The data of 37 participants were used in the analyses. Five participants' data were discarded due to high (>15%) error rates.

The data were first analyzed in terms of an overall lexical effect and an overall phonotactic effect. Mean RTs and differences between RTs are displayed in the top third of Table 3. There was a 64 ms difference between words and nonwords in the expected direction, which was significant by subjects, $F(1,36)=10.74$, $p=.002$, and by items, $F(1,47)=8.70$, $p=.005$. There was also a 31 ms phonotactic effect in the expected direction, which was marginally significant by subjects, $F(1,36)=2.76$, $p=.105$, and by items, $F(1,47)=2.24$, $p=.141$. Twenty-two of the 37 participants showed the phonotactic effect in the expected direction. The interaction was not significant by subjects or by items, $F<1$. Errors to target items were very infrequent, ranging from less than 1% in the illegal-word condition to 2.2% in the legal-nonword condition.

The data in Table 3 show that participants were fastest in the illegal-word condition, and were, on average, 94 ms slower in the legal-nonword condition. In the word conditions, the phonotactic effect was weak (17 ms) and nonsignificant, $F<1$ by subjects and by items. The phonotactic effect was larger (45 ms) for the nonword condition. This was not significant by subjects, $F(1,36)=1.63$, $p=.21$, and was marginally significant by items, $F(1,47)=2.14$, $p=.15$.

2.1.3 Discussion

The purpose of this experiment was to examine the interplay of phonotactic and lexical cues to segmentation. Both lexical and phonotactic cues were found to have important influences on word segmentation. Certain predictions were confirmed. Overall, word RTs were significantly faster than nonword RTs. In addition, there was a marginally significant phonotactic effect, such that RTs in the phonotactically illegal continuation were faster than RTs in the phonotactically legal continuation. More specifically, the illegal-word condition, which has the advantage of involving both phonotactic and lexical information, was the fastest condition overall, while the legal-word condition, which has neither advantage, was slowest.

So what do these findings suggest? They suggest that the listener has various resources from which he draws when listening to connected speech. Lexical information is a crucial resource, as shown by the large (64 ms) overall lexical effect. However, as the speech signal also contains phonotactic cues, the listener can rely on this additional information source. The overall phonotactic effect was small (31 ms) in comparison to the lexical effect, and was marginally significant, but still trended in the expected direction. This suggests that phonotactics appeared to play a role, albeit a small one, in aiding segmentation. Phonotactic information appeared to be more useful when lexical information was ambiguous (i.e., the CW was a nonword) than when lexical information unambiguously specified the boundary (i.e., the CW was a word). The phonotactic effect for words (17 ms) was not significant, but the phonotactic effect for nonwords trended towards significance (45 ms).

How do these effect sizes compare to those found in Experiment 3, the experiment with which it is most directly related? When comparing the data from the stop/nasal subset of stimuli (e.g., aligned ‘sib case’ vs nonaligned ‘sis case’) to the nonword subset of the present results (e.g., legal-nonword ‘var-case’ vs illegal-nonword ‘zan-case’), the phonotactic effect is found to be nearly equivalent (52 ms vs 45 ms). This replication strengthens the claim that phonotactic effects influence between-word segmentation. Because lexical effects were not examined in Experiment 2b, no lexical comparison can be made.

One could argue that experimental setup does not allow the researcher to know exactly what the listener perceived the CW to be. Because predictions are being made about RTs based on HOW the listener perceived the CW (i.e., as a word or a nonword), it is important to find out how they heard it. Given that the trends in the data are in line with the predictions (i.e., words allow for easier segmentation, and thus, faster RTs than nonwords), it is likely that listeners are hearing the CW as intended. However, this needs to be confirmed. Therefore, Experiment 4 was rerun with the addition of a secondary multiple-choice task aimed at gathering information regarding listeners’ perception of the CW.

2.2 Experiment 5--Replication of Experiment 4 with a multiple-choice task

2.2.1 Method

Subjects. Thirty-eight students from the same pool as the previous experiments participated for class credit.

Stimuli and Procedure. The stimuli and methods are identical to those of Experiment 4, with the addition of the multiple-choice questions and the ensuing instructional changes.

For each target item, four multiple-choices were created. Each choice, labeled A-D, was phonetically similar (or identical to) the CW. The choices always consisted of at least one word and one nonword. The multiple-choice question only appeared after target trials. The multiple-choice items appeared on the computer screen after participants made the lexical decision response.

Participants were instructed on the multiple-choice task once it was clear that they understood the lexical-decision task. Participants were told that after they made the word-nonword response, a set of four items, labeled A-D, would appear on the computer screen. They were told to pick the item that matched the second word of the phrase that they had just heard, and to write the corresponding letter (A-D) of the correct answer on the answer sheet which was then presented to them. They were told outright that the purpose of the task was to ensure that they were listening to the entire phrase, and that because it was meant to 'keep them on their toes,' they would only be given the multiple-choice task on about one fourth of the trials.

2.2.2 Results

The data of 28 participants were used in the analyses. Ten participants' data were discarded because of high (>15%) error rates. Across all discarded subjects' data, the majority (61%) of their errors were due to a failure to respond. Most likely, this was due to the addition of the multiple-choice task, as only five participants' data were discarded in Experiment 4.

The first set of analyses excluded only incorrect responses (or no responses) to the lexical decision task. Responses were not discarded on the basis of the multiple-choice responses. Mean RTs and differences between RTs are shown in the middle section of Table 3. There was a main effect of lexical status, with a 66 ms RT advantage for words over nonwords. A repeated measures ANOVA by subjects and by items revealed this difference was significant, $F(1,27)=11.84$, $p=.02$, $F(1,46)=15.25$, $p=.01^3$. There was a smaller phonotactic effect, with the illegal continuation stimuli 34 ms faster than legal-continuation stimuli. This was significant by subjects, $F(1,27)=4.76$, $p=.03$, but not by items, $F(1,46)=1.15$, $p=.30$. A closer analysis of the individual items revealed that some of the stimuli in the illegal-word condition were responded to particularly slowly, thus creating a higher mean RT for this condition. As a result, the difference between the illegal-word and the legal-word conditions (i.e., the phonotactic effect) was smaller, contributing to the nonsignificant results in the item analysis. There was no phonotactic x lexical interaction by subjects or by items, $F<1$. The size of the phonotactic effect was

essentially the same across both words (32 ms) and nonwords (37 ms). Error rates to target items were low across all four conditions, ranging from less than 1% in the illegal-word condition to 2.4% in the legal-nonword condition.

Finally, the multiple-choice responses were evaluated. Most participants performed particularly poorly on this task. These results were quite unexpected, as post-experiment questionnaires revealed that participants expressed no concerns or difficulties with the task or with the stimuli. On a scale of 1 (very easy) to 10 (very difficult), the mean rating was 4.08. The mean number of errors (out of 48) was 15.54 (32%).

An analysis of the multiple choice errors indicated that over half (53%) of all the errors involved mishearings of the same lexical status: 26% of the total number of errors involved misperceiving one word as another word (e.g., ‘class mind’ heard as ‘glass mind’); 27% involved misperceiving one nonword as another nonword (e.g., ‘vab test’ heard as ‘vap test’).

Forty-seven percent of the total number of errors were misperceptions of lexical status (i.e., hearing a nonword as a word, e.g., ‘bess food’ as ‘bass food’ or hearing a nonword as a word, e.g., ‘harm food’ as ‘sharm food’). Because this latter type of misperception could influence a participant’s RTs (based on predictions that a CW perceived as a word would result in an easier segmentation, and thus a faster RT to the TW, or conversely, a CW perceived as a nonword would result in a more difficult segmentation, and thus a slower RT to the TW), lexical decision responses corresponding to these errors were removed, and the ANOVA was re-run. The results remained unchanged. There was again a main effect of lexical status, with 47 ms RT advantage for

words over nonwords. This was significant by subjects, $F(1,27)=7.57$, $p=.01$ and marginal by items, $F(1,46)=3.34$, $p=.07$. There was also a main effect of phonotactic status by subjects, with the illegal continuation stimuli 31 ms faster than the legal continuation stimuli, $F(1,27)=4.17$, $p=.051$. This effect was nonsignificant, however, by items, $F(1,46)=1.89$, $p=.17$. There was no interaction either by subjects or by items, $F(1,27)<1$.

2.2.3 Discussion

Experiment 5 was designed with the intent of finding out how listeners perceived the CW. Their perception of the CW (i.e., as a word or a nonword) could influence how they responded to the TW. Despite participants poor performance on this secondary task, an analysis of the data that included incorrect responses to the multiple-choice task and an analysis of the data that excluded incorrect responses to the multiple-choice task show the same results. There was an effect of lexical status, with words allowing for easier segmentation than nonwords. There was also a slightly smaller effect of phonotactics, with phonemes forming an illegal continuation at the boundary allowing for easier segmentation than phonemes forming a legal continuation at the boundary. These data essentially replicate the findings of Experiment 4 and thus demonstrate the combined contribution of lexical and phonotactic information in between-word segmentation.

The data up to this point suggest that listeners can and do make use of multiple sources of information when segmenting between words. Clearly, lexical information aids listeners enormously, as demonstrated by the significant lexical effects found in both

Experiment 4 and Experiment 5. Phonotactic information also appears to be used in segmentation, as a small effect has been consistently found across Experiments 4 and 5. To some extent, the phonotactic effect appears to be influenced by the presence or absence of lexical information. When phonotactics serves as the primary source of information in between-word segmentation (Experiment 3), or when phonotactics is present in conjunction with ambiguous lexical information (the nonword condition in Experiment 4 and 5), the effect is larger than when phonotactic information is present in conjunction with unambiguous lexical information (the word condition in Experiment 4 and 5). Although the trends suggest that lexical information mediates phonotactic effects, none of the lexical x phonotactic interactions were significant in either Experiment 4 or Experiment 5.

2.3 Experiment 6--Testing the limits of the phonotactic effect

If the presence of lexical segmentation cues results in a small phonotactic effect, then perhaps increasing these lexical segmentation cues may make the phonotactic effect disappear. What if a lexical candidate is so highly activated as to allow the listener to recognize the CW even before its acoustic offset, thus making its boundaries known before the TW is spoken? Will a phonotactic effect still emerge, or will it disappear completely?

To test this idea, participants were presented with three-word utterances in which the CW was a multi-syllabic (3+ syllables) word. The CWs became lexically unique, and thus maximally activated by the offset of the second syllable. Because of the maximal activation level of the CW, it was predicted that its offset would be anticipated, and thus segmentation would easily be achieved without the need for phonotactic information. If this is the case, then the phonotactic effect should be negligible across the legal- and illegal-continuation stimuli in the word condition.

In addition to presenting listeners with multi-syllabic words, multi-syllabic nonwords were also used. These stimuli were used to test another obvious prediction, namely, whether the lack of reliable lexical information would result in an increased phonotactic effect. Lexical information in nonwords (long or short) is ambiguous--there may be lexical activation of a few different candidates, but no single candidate in particular. And at the point in the signal in which the input deviates from all lexical candidates (i.e., the deviation point), there should be only residual activation of a few

similar sounding lexical candidates. With little lexical information to go on, will phonotactic information become more important in aiding segmentation? Experiment 6 was designed to explore these ideas.

2.3.1 Method

Participants. Forty-nine students from the same pool as the previous experiments participated for class credit.

Stimuli. Forty-eight sets of stimuli were chosen. As with Experiment 4 and 5, the stimuli in each set consisted of Word0 followed by the CW and the TW. The TWs were the same as those used in the previous experiments. The CWs were three- or four-syllable words and nonwords. The nonwords were created by changing three or four phonemes in a word (e.g., cellular --> shalluber). Just as in the previous experiments, there were four conditions: illegal-word ('nice figurine case'), legal-word ('nice cellular case'), illegal-nonword ('nice mipyurveen case'), legal-nonword ('nice shalluber case'). The stimuli can be found in Appendix C. As before, within each phonotactic condition, the offset of the CW and the onset of the TW were the same. They were recorded in the same manner as the previous experiments. One nagging concern of Experiments 4 and 5 was that the environment at the boundary between the CW and TW within each phonotactic condition was not acoustically identical. To alleviate this concern, the boundary consisting of the offset of the CW and the onset of the TW within each phonotactic condition were made identical by cutting off the offset of the CW plus the TW (e.g., /nk/ in 'figurine case') and splicing it onto the other CW (e.g., 'mipyurveen

case'). The splicing was done at a point that included *at least* part of the final phoneme of the CW, if not more. Care was taken to ensure that the splicing was not noticeable. This was done within each phonotactic condition for all 48 sets of stimuli.

Because the 152 filler stimuli used in Experiments 4 and 5 were varied along many dimensions, including CW length, they were able to be used in this experiment. Ten new practice trials were created to replace some of the practice trials containing short CWs.

Procedure. The procedure was identical to Experiment 5.

2.3.2 Results

The data of 36 participants were used. The data were first analyzed for a main effect of lexical status and phonotactics. Mean RTs and differences between RTs are displayed in the bottom third of Table 3. Overall, word RTs were 80 ms faster than nonwords, an effect that was highly significant by subjects, $F(1,35)=17.12, p=.000$ and by items, $F(2,47)=11.53, p=.001$. There was no main effect of phonotactics, $F<1$ by subjects and by items.

Error rates across the target items were higher for this experiment. Errors were low in the word conditions (less than 1% for legal continuation words and 2.5% for illegal continuation words) and were higher for nonwords (6.1% for legal continuation nonwords and 7.3% for illegal continuation nonwords). This lexical effect was significant by subjects and by items, $F(1,35)=31.26, p=.00, F(1,47), F(2,47)=14.31,$

$p=.00$, respectively. The phonotactic effect was marginal by subjects, $F(1,35)=2.41$, $p=.13$, and nonsignificant by items, $F(1,47)=1.0$, $p=.32$. Neither the interaction by subjects nor by items was significant, $F<1$.

The data show a small, nonsignificant phonotactic effect (26 ms) in the word condition. The phonotactic effect in the nonword condition was in the opposite direction (i.e., legal-continuation stimuli were responded to on average 27 ms faster than illegal-continuation stimuli), but this effect was not significant. Interestingly, the lexical effect was twice as large in the illegal-continuation stimuli (107 ms) than in the legal-continuation stimuli (54 ms). This difference was not significant by subjects or by items, however, $F_1(1,35)=1.83$, $p=.185$, $F_2(1,47)=1.95$, $p=.169$.

2.3.3 Discussion

The data from Experiment 6 revealed some interesting findings. The large lexical effect, which had been found in all the previous experiments was found again, both in the RT data and in the error data. More importantly, though, there was a reduced phonotactic effect. In the word condition, the phonotactic effect was not significant. This makes sense in light of the large lexical effect created by the multi-syllabic words. All the multi-syllabic words in this experiment became unique well before their offsets, thus making them highly activated and, according to some theories (e.g., Cohort Theory: Marslen-Wilson & Welsh, 1978), even recognized prior to their offset. Their offsets were highly predictable, and thus, may have rendered phonotactic information not useful. These data

thus confirm the prediction set forth earlier: The presence of unambiguous lexical information outweighs information provided by phonotactic segmentation cues.

Surprising data were obtained in the nonword condition. Instead of the illegal nonword condition being significantly faster than the legal nonword condition, which would have indicated an increased role of phonotactics, there was no effect of phonotactics. The illegal nonword condition was actually slower than the legal nonword condition, but not significantly so.

How can the lack of a phonotactic effect in the nonword stimuli be explained? It is suspected that the difficulty in processing multi-syllabic nonwords may have contributed to the lack of an effect. Phonotactic constraints within a lengthy multi-syllabic item (word or nonword) may signal multiple possible segmentation points, thus making segmentation based solely on phonotactics difficult. With words, reliance on phonotactics is not necessary because lexical information can guide segmentation.

CHAPTER 3

GENERAL DISCUSSION

The goal of this paper was to better understand the contribution of phonotactics in word segmentation. Previous research on phonotactics (McQueen, 1998; Norris et. al., 1997; van der Lugt, 2001) suggests that it is a reliable source of information in the segmentation of syllable boundaries within isolated words. The research is unclear, however, on whether the influence of phonotactics extends to segmentation between words, where acoustic/phonetic cues to word boundaries are already present. Furthermore, little research has been done to investigate how phonotactic information combines with other sources, such as lexical information, to aid in word segmentation. The experiments presented here attempt to shed light on these issues.

Part I: Replicating phonotactic effects in between-word segmentation

The experiments in Part I replicate and extend McQueen's findings that phonotactic information does aid in segmentation. The first two experiments (Experiments 1 and 2) showed that listeners used phonotactic cues to locate word-internal boundaries in English. One interesting finding was that the size of this effect varied as a function of the phonetic makeup of the stimulus. RTs were slower and there was a much

larger phonotactic effect when the onset of the embedded word was a liquid consonant than when the onset was a stop consonant. Experiment 3 extended the word-spotting findings to two-word utterances. The stimuli of Experiment 2 were re-recorded as two-word utterances (e.g., 'sibcase' → 'sib case'), which allowed for the naturally-occurring presence of acoustic/phonetic information produced when the items were spoken as two words rather than as a single word. The results of this experiment demonstrated once again that phonotactics contributed in segmentation. In the stop/nasal-onset stimuli the effect was significant. This time, however, the liquid-onset stimuli did not show the effect in the expected direction.

Why was the phonotactic effect different for the stop-onset stimuli than for the liquid-onset stimuli across all three experiments? A likely possibility has to do with acoustic/phonetic differences among the stimuli. The presence of acoustic/phonetic information (i.e., silence) in the stop-onsets to aid in locating a word boundary can not only explain the smaller phonotactic effect found for stop-onsets in Experiments 1 and 2, but also the faster overall RTs for these items. Conversely, the lack of a salient acoustic/phonetic segmentation cue can explain why the phonotactic effect was larger for liquid-onsets in the word-internal boundary condition (Experiments 1 and 2). That the phonotactic effect was in the opposite direction in the between-word boundary condition (Experiment 3) was not expected, however. Previous research (Nakatani & Dukes, 1977) suggests that allophonic variation in /l/ and /r/ is an important cue to signal a word onset. Just as silence served as a cue in the stop-onset stimuli, if this acoustic/phonetic

information served as a cue to signal a word boundary, the result could be a smaller phonotactic effect for these items. Why the effect would be in the opposite direction, however, is not entirely clear.

What do the findings of Part I say about the effect of phonotactic information on the segmentation of within-word and between-word boundaries? Two important observations were made. First, it appears that the size of the phonotactic effect may be influenced by the presence or absence of acoustic/phonetic cues. The evidence presented in Experiments 1-3 indirectly support this contention, as acoustic/phonetic cues were not intentionally manipulated. The data of Vroomen et. al. (1998) do support this observation, however, as they experimentally demonstrated that phonological information (i.e., word stress in Finnish) can influence the effects of acoustic/phonetic information (i.e., vowel harmony) on segmentation.

The second observation was the realization that there are important differences between segmenting within words and segmenting in context (i.e., between words). Words in context inherently have additional acoustic/phonetic information that may be used by the listener in segmentation. Experiment 3, which looked at words in context, demonstrated that cues to signal word boundaries differ as a function of the acoustic/phonetic makeup of the stimuli. The acoustic/phonetic cues that signal stop-onsets are different from those that signal liquid-onsets. The data from Experiment 3 showed that stop-onsets in context behaved differently than liquid-onsets in context. In addition, a comparison of the phonotactic effects in the liquid-onset stimuli from

Experiment 2 (within-word segmentation) with the liquid-onset stimuli from Experiment 3 (between-word segmentation) shows that the effects differ, and suggest that acoustic/phonetic differences may be the cause.

That there are acoustic markings to signal word boundaries is not surprising. Gow and Gordon (1995) showed that acoustic markings can allow listeners to differentiate between syllables that form the onset of a word (e.g., 'lips' in 'two lips')and syllables that do not form the onset of a word ('lips' in 'tulips'). Davis et. al. (2002) showed that acoustic differences in word-embedded syllables (e.g., 'cap' in 'captain') help listeners discriminate short words ('cap') from the those that form the beginnings of longer words ('captain').

So what are the implications of these findings on McQueen's (1998) data? McQueen's data are important for demonstrating experimentally that phonotactic information can be used in locating within-word boundaries. The narrow subset of stimuli he used, however, hid the critical finding that phonotactic effects can be influenced by the phonetic makeup of the embedded words. One must consider the presence of acoustic/phonetic cues that can constrain the phonotactic effect. Additionally, given that acoustic/phonetic cues that signal word onsets may even be more salient than acoustic/phonetic cues that signal syllable onsets, it is important to constrain the interpretation of McQueen's findings to the stimuli being used--liquid-onset embedded words, and to the issue being studied--within-word segmentation.

In sum, the results of the experiments in Part I show that phonotactic information does aid the listener in locating within-word and between-word boundaries. While the

size of the phonotactic effect varied considerably across different stimulus types, the overall consistency in which the phonotactic effect emerged is solid evidence of its use in segmentation.

Part II: Understanding the contribution of phonotactic and lexical information in between-word segmentation

The experiments in Part II investigated whether phonotactics would continue to play a significant role in segmentation when other sources (i.e., lexical information) were also available to listeners. Given that the phonotactic effect appeared to be influenced by the acoustic/phonetic makeup of the stimuli in Part I, it was hypothesized that the size of the phonotactic effect would be influenced by the presence of lexical information too.

Experiments 4-6 consisted of stimuli of varying combinations of phonotactic and lexical information. The results across all three experiments always showed a large lexical effect. Listeners were faster at responding to the TW when the CW was a word than when it was a nonword. This makes sense. When the item is a nonword, theoretically, the listener never knows where the offset of the nonword is until he or she recognizes the word immediately following. Therefore, participants will be slower at locating the word boundary between the CW and the TW, resulting in slower RTs to identifying the TW.

In addition to the lexical effect, Experiments 4 and 5 also showed a consistently small effect of phonotactics. To some extent, the phonotactic effect appeared to be influenced by the lexical status of the stimulus. The phonotactic effect was smaller in the

word condition when compared to the nonword condition for both experiments, but this was never statistically significant. Figure 1 graphically displays the size of the phonotactic effect as a function of lexicality across Experiments 4-6, as well as in Experiment 3 (which only had a nonword condition). The trend in the data suggests that when lexical information is present to help in identifying word boundaries, phonotactic information is not as useful. When lexical information is unable to specify a word's boundary (as in the case of nonwords), phonotactic information becomes more important.

Given the possibility that lexical effects may mediate phonotactic effects, as suggested by the data of Experiments 4 and 5, Experiment 6 was designed. The goal of Experiment 6 was to see whether phonotactic information would still be used when lexical information unambiguously specified the offset of the CW (as with the case of multi-syllabic words whose uniqueness point fell well before the word's offset), and conversely, whether phonotactic information would become more useful when ambiguous lexical information made the offset of the CW completely unknown (as with the case of multi-syllabic nonwords). The results again showed a large lexical effect. This time, however, the phonotactic effect was not significant, nor was the interaction. The nonsignificant phonotactic effect in the word condition suggests that listeners were able to recognize the word before its acoustic offset, eliminating the need for phonotactic information to assist in segmentation. To find a significant phonotactic effect in the nonword condition would have strengthened the argument that phonotactic effects are influenced by lexical effects. However, this was not found. The lack of an effect in the nonword condition is attributed to the difficulty in processing the stimuli.

What can be concluded from the experiments of Part II regarding the influence of phonotactic information in conjunction with lexical information? First and foremost, the results suggest that lexical information has a much greater influence on segmentation than does phonotactics. All six experiments showed large and reliable lexical effects (average effect=70 ms) . Across both phonotactic conditions, words were responded to more quickly than nonwords in all three experiments. This finding lends support to a lexically-based model of segmentation, which considers segmentation a consequence of successful word recognition. Second, the small but consistent effects of phonotactics that were found across experiments 4 and 5 suggest that phonotactic information can be used by listeners in segmenting speech. When phonotactic information requires a boundary be placed between two co-occurring phonemes, listeners use this information as a cue to signal a word boundary. That the effect of phonotactics was small, though, suggests that listeners do not rely heavily on it as a source of information.

What about the combined influence of phonotactic and lexical information on word segmentation? The overall pattern of data in Experiments 4-6 suggest that lexical information plays the primary role in segmentation, and that the influence of phonotactics is secondary. In Experiments 4-6 , RTs were always faster in the word conditions (legal and illegal) than the nonword conditions (legal and illegal). The influence of phonotactics could also be seen, though. Within each lexical condition, the illegal-continuation stimuli were responded to more quickly than legal-continuation stimuli (with the exception of the nonword condition of Experiment 6). More specifically, RTs in the illegal-word condition were always fastest, and RTs in the legal-nonword condition were

slowest (excepting Experiment 6). There was a trend of an interaction in Experiment 4 between the two sources of information, such that the phonotactic effect was larger for words than for nonwords, but this never reached significance. The trend, which was specifically tested, and in part confirmed in the word conditions of Experiment 6, suggested that lexical information can mediate the effects of phonotactics. More specifically, in Experiment 6, the phonotactic effect diminished in the presence of a very strong lexical cue.

That multiple cues can combine, and even interact, to influence segmentation is supported by others' research. Norris et al (1995) found that effects of metrical stress were modulated by lexical competition effects. The more lexical competitors that are activated, the greater the size of the prosodic effect. Similar effects were found by Vroomen & de Gelder (1995). Vroomen et. al. (1998) focused on how listeners deal with word stress and vowel harmony (in Dutch) as segmentation cues and found that the cues were used in an interdependent way. Stress was the strongest cue, and its presence reduced the contribution of vowel harmony. Christensen et. al. (1998) looked at the contributions of multiple cues using a simple recurrent network. The network was provided with information about phonemes, relative lexical stress, and boundaries between utterances. On their own, these segmentation cues proved unreliable; however, after training, the conjunction of these cues allowed the model to reliably find word boundaries.

In sum, the data suggest that phonotactic information does work in conjunction with lexical information to influence segmentation, but to a smaller degree, and that the effects of phonotactic information may diminish in the presence of a maximally activated lexical item.

Many of the most frequently cited models of word recognition (e.g., TRACE, Shortlist) are lexically-based models that use a competition process to achieve successful recognition (and thus segmentation). Current instantiations of these models do not implement a separate segmentation mechanism to allow for additional information sources (e.g., acoustic/phonetic, phonotactic) to have an input on constraining implausible candidates or boosting activation levels of plausible candidates (or both). If these models are to effectively capture human behavior, it is necessary that they incorporate some process or mechanism to explain the phonotactic effects found in the experiments presented in this paper. This would not be difficult to do. A lexical model could still continuously activate candidates at any point at the speech input; the addition of a phonotactic segmentation process could then be engaged to influence the activation levels of those candidates. In a scenario such as this, lexical information still has the primary role, but allows for phonotactic information to have some influence, especially when lexical information is ambiguous or lacking. Norris et. al. (1995) propose just such an implementation using Shortlist and the Metrical Segmentation Strategy heuristic. Gow and Gordon (1995) also propose such an implementation in their Good Start model using acoustic/phonetic information (e.g., segmental durations, full vowels).

The research presented here shows that not all cues are equally influential, and also implies that cues can interact in interesting ways. Future research needs to be done to more fully understand how cues combine to influence word segmentation.

Experiments investigating the influence of high-level cues, such as semantic information, in biasing the activation levels of lexical items could be performed to further test the limits of phonotactic effects. In addition, experiments pitting conflicting segmentation cues could help determine which sources of information dominate in segmentation.

In closing, this study demonstrates the importance of investigating the combined influence of segmentation cues in word recognition. In and of themselves, phonotactic and lexical cues to segmentation have been shown to be useful. However, their combined influence reveals a complexity that would never have been uncovered had the cues been studied in isolation. If researchers are to obtain a realistic view of how listeners segment speech, not only do they need to consider all sources of information available to listeners, but also how these sources combine.

Endnotes

¹The set of five nasal-onset stimuli were analyzed separately to see whether they showed the same phonotactic effect as did the stop-onsets, with which they were grouped. The results showed that the nasal-onsets behaved more similarly to the liquid-onsets in that they produced a large (223 ms) phonotactic effect. When removed from the stop-onset stimuli, the overall phonotactic effect was reduced but was still marginally significant, $F(1,41)=2.72$, $p=.107$. There was again no interaction of alignment and stimulus type, $F(1,41)=1.52$, $p=.224$. Removal of the nasal-onset stimuli in Experiments 2 and 3 does not change the results obtained with their inclusion, and therefore are not reported.

²A rough estimate of the F0 contour was calculated by finding the peak F0 and valley F0 measurements for each syllable (or word), subtracting one from the other, and comparing the resulting differences across experiments. No significant differences would suggest that the prosodic contour was similar across the two conditions (e.g., ‘sibcase’ vs ‘sib case’).

³ANOVA by items was performed on 47 of the 48 sets of stimuli. One set of items was excluded because one stimulus within the set was inadvertently left out in a list.

Experiment	Example Stimulus
1 & 2	aligned: sibcase nonaligned: siscase
3	aligned: sib case nonaligned: sis case
4 & 5	illegal continuation word: nice car case nonword: nice var case legal continuation word: nice tan case nonword: nice zan case
6	illegal continuation word: nice figurine case nonword: nice mipurveen case legal continuation word: nice cellular case nonword: nice shalluber case

Table 1. Examples of stimuli used in Experiments 1-6.

Experiment	Alignment Type		Difference
	aligned	nonaligned	
1 (N=24)	655 (325)	736 (313)	82 ms
2 (N=34)	1091 (274)	1169 (315)	78 ms
3 (N=38)	1040 (212)	1041 (185)	1 ms

Table 2. Mean RTs for Experiments 1-3 (Part I) as a function of alignment (sd in parentheses). Mean differences between alignment conditions are also shown.

Experiment	Lexical Status of CW	Phonotactic Status of CW		Difference
		legal	illegal	
4 (N=37)	word	1026 (219)	1009 (219)	17 ms
	nonword	1104 (284)	1059 (250)	45 ms
	Difference	78 ms	50 ms	
5 (N=28)	word	964 (175)	932 (176)	32 ms
	nonword	1033 (189)	996 (177)	37 ms
	Difference	69 ms	64 ms	
6 (N=36)	word	969 (211)	943 (215)	26 ms
	nonword	1023 (236)	1050 (269)	-27 ms
	Difference	54 ms	107 ms	

Table 3. Mean RTs for Experiments 4-6 (Part II) as a function of lexical status and phonotactic status (sd in parentheses). Mean differences between phonotactic conditions and lexical conditions are shown in the margins.

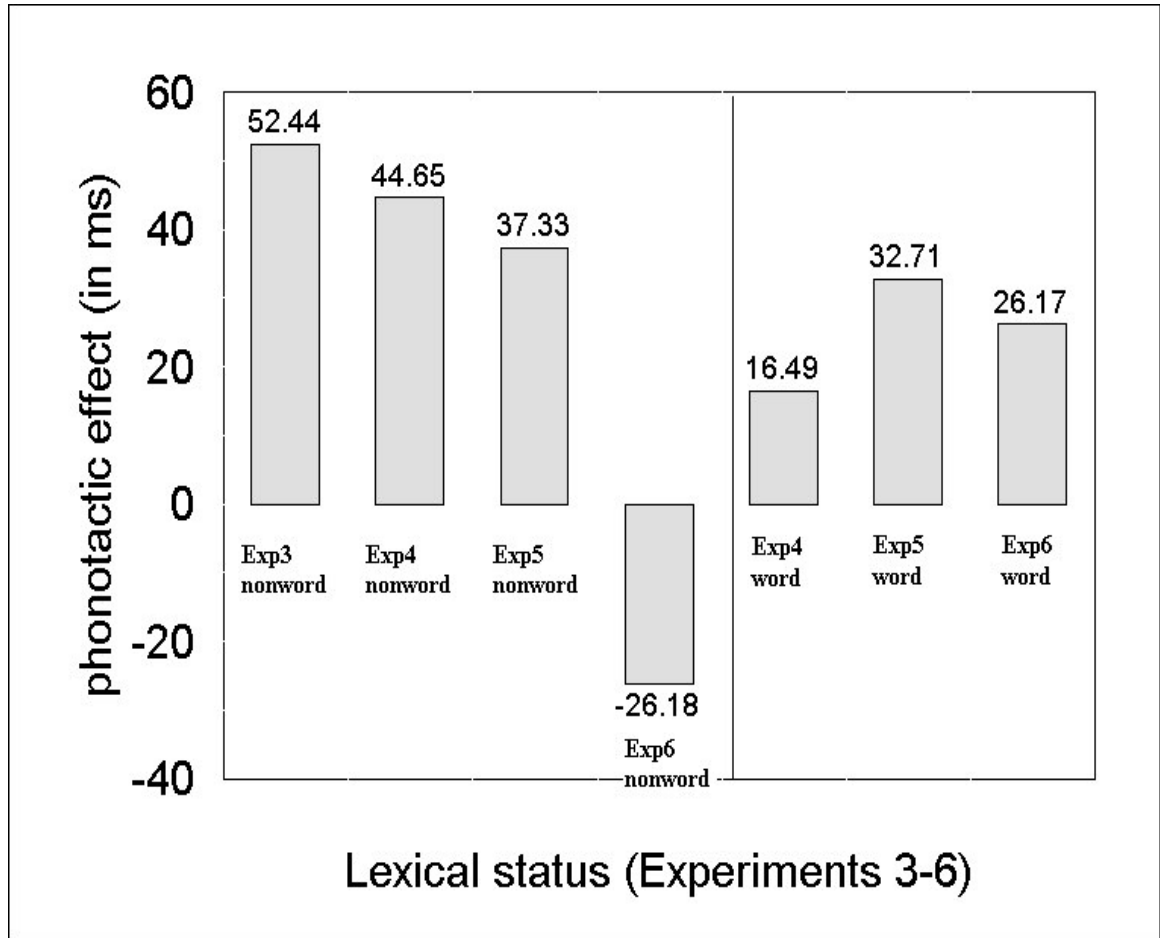


Figure 1. Phonotactic effect across experiments as a function of lexical status.

LIST OF REFERENCES

Christiansen, M.H., Allen, J., & Seidenberg, M.S. (1998). Learning to segment speech using multiple cues: A connectionist model. Language and Cognitive Processes, 13(2/3), 221-268.

Church, K.W. (1987). Phonological parsing and lexical retrieval. Cognition, 25, 53-69.

Cole, R.A. & Jakimik, J. (1980). A model of speech perception. In R. A. Cole (Ed.), Strategies of Information Processing (pp.67-116). London: Academic Press.

Cowan, N. (1991). Recurrent speech patterns as cues to the segmentation of multisyllabic sequences. Acta Psychologica, 77, 121-135.

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

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

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. Journal of Experimental Psychology: Human Perception and Performance, 14(1), 113-121.

Davis, M.H., Marslen-Wilson, W.D., & Gaskell, M.G. (2002). Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. Journal of Experimental Psychology: Human Perception and Performance, 28(1), 218-244.

Goldinger, S.D., Luce, P.A., Pisoni, D.B., & Marcario, J.K.(1992). Form-based priming in spoken word recognition: The roles of competition and bias. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 1211-1238.

Gow, D.W., & Gordon, P.C. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. Journal of Experimental Psychology: Human Perception and Performance, 21(2), 344-359.

Hayes, J.R., & Clark, H.H. (1970). Experiments in the segmentation of an artificial speech analog. In J.R. Hayes (Ed.), Cognition and the development of language (pp. 221-233). New York: Wiley.

Johnson, E.K., & Jusczyk, P.W. (2001). Word segmentation by 8-month olds: When speech cues count more than statistics. Journal of Memory and Language, 44, 548-567.

Jusczyk, P.W., & Houston, D.M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. Cognitive Psychology, 39, 159-207.

Lehiste, I. (1960). An acoustic-phonetic study of internal open juncture. Phonetica, 5, 1-54.

Luce, P.A. (1986). A computational analysis of uniqueness points in auditory word recognition. Perception and Psychophysics, 39, 155-158.

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

Massaro, D.W., & Cohen, M.M. (1983). Phonological constraints in speech perception. Perception and Psychophysics, 34, 338-348.

McClelland, J.L., & Elman, J.L. (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.

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

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

Norris, D., McQueen, J.M., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21(5), 1209-1228.

Norris, D., McQueen, J.M., Cutler, A., & Butterfield, S. (1997). The Possible-Word Constraint in the segmentation of continuous speech. Cognitive Psychology, *34*, 191-243.

Onishi, K.H., Chambers, K.E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. Cognition, *83*, B13-B23.

Pitt, M.A., & McQueen, J.M. (1998). Is competition for coarticulation mediated by the lexicon? Journal of Memory and Language, *39*, 347-370.

Pitt, M.A., & Shoaf, L.C. (In press). Revisiting bias effects in word-initial phonological priming.

Pollack & Pickett (1964). Intelligibility of excerpts from fluent speech: Auditory vs structural context. Journal of Verbal Learning and Verbal Behavior, *3*, 79-84.

Radeau, M., Morais, J., Mousty, P., & Bertelson, P. (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.

Saffran, J.R., Newport, E.L., & Aslin, R.N. (1996). Word segmentation: The role of distributional cues. Journal of Memory and Language, *35*, 606-621.

Shillcock, R. (1990). Lexical hypotheses in continuous speech. In G.T.M. Altmann (Ed.) Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives (pp. 24-49). Cambridge, MA: MIT Press.

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

van der Lugt, A.H. (2001). The use of sequential probabilities in the segmentation of speech. Perception and Psychophysics, *63*(5), 811-823.

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

Vroomen, J., van Zon, M., & de Gelder, B. (1996). Cues to speech segmentation: Evidence from juncture misperceptions and word spotting. Memory and Cognition, *24*(6), 744-755.

Vroomen, J., Tuomainen, J., & de Gelder, B. (1998). The roles of word stress and vowel harmony in speech segmentation. Journal of Memory and Language, 38, 133-149.

Wurm, L.H. (1997). Auditory processing of prefixed English words is both continuous and decompositional. Journal of Memory and Language, 37, 438-461.

Zwitzerlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. Cognition, 32, 25-64.

APPENDIX A
STIMULI USED IN PART I^a

I. Stop/Nasal-Onset Stimuli

Stimulus	Nonaligned	Aligned
camp	teskæmp	tetʃkæmp
case	sɪskes	sɪbkes
cast	bɪskæst	bɪfkæst
coast	vuskost	vudkost
cost	fɪskøst	fɪdkøst
crowd	tʃæskrawd	tʃæfkrawd
cut	fuskʌt	fubkʌt
kiss	tæskɪs	tæfkɪs
milk	væsmɪlk	vægmɪlk
mind	rusmaɪnd	rupmaɪnd
mouth	taɪsmawθ	taɪgmawθ
name	vɪjsnem	vɪjgnem
noise	pɒsnɔɪz	pɒgnɔɪz
page	ŋɛspedʒ	nejbpedʒ
pass	rawspæs	rawgpæs
piece	næspɪs	næspɪs
plane	næsplen	nætʃplen
play	dʒɪsple	dʒɪdple
plaque	θɪsplæk	θɪgplæk
plate	vɪsplet	vɪdʒplet
price	fæsprais	fædʒprais
pride	fæspraid	fædʒpraid
test	fɪstest	fɪptest
toast	ʃɪstost	ʃɪgtost

Appendix A (continued)

II. Liquid-Onset Stimuli

Stimulus	Nonaligned	Aligned
rake	tʃaɪkrek	tʃaɪnrek
race	pægrais	pævrais
road	gɪkrod	ɡɪmrod
rain	vopren	volren
rip	wɪbrɪp	wɪdʒrɪp
rag	wɛgræg	wɛsræg
rat	lɔɪkræt	lɔɪvræt
rope	futrop	fuzrop
rose	sæbroz	sætʃroz
red	ʃegred	ʃetʃred
room	dæfrum	dædʒrum
roof	ʃɪfruf	ʃɪmruv
lake	θuklek	θuvlek
lace	θegles	θetʃles
leg	zæblɛɡ	zædʒlɛɡ
life	dɪklaɪf	dɪdʒlaɪf
lime	rɪflaɪm	rɪnlɪaɪm
laugh	ʃeɪflæf	ʃeɪnlæf
load	veklod	vezlɒd
loan	tæslon	tævlon
leash	mawblɪʃ	mawtʃlɪʃ

^aFor Experiments 1 and 2. For Experiment 3, the same stimuli were used, but they were recorded as two-word utterances (e.g., teskæmp → tes kæmp)

APPENDIX B

STIMULI USED IN PART II (EXPERIMENTS 4 AND 5)

Word0	Critical Words				Target Word
	Word- Illegal	Nonword- Illegal	Word- Legal	Nonword- Legal	
rich	ʃɪp	θɪp	geŋ	fɾeŋ	camp
nice	tæn	zæn	kɑr	vɑr	case
will	tʃɪp	plɪp	slɪŋ	dʒɪŋ	cast
did/wide	dʌmp	jʌmp	lɔŋ	zɔŋ	coast
high	kræb	mæb	bræs	dæs	cost
huge	ʃɑp	dʒɑp	jʌŋ	ʃʌŋ	crowd
bad	taɪp	maɪp	tʌŋ	grʌŋ	cut
can	kawnt	bawnt	ʃaɪn	gaɪn	dance
late	kɔrt	zɔrt	tɪn	ʒɪn	date
his	wɪk	fɪk	gʌn	gæn	deal
may	tʌtʃ	nʌtʃ	fæn	smæn	dirt
cant/want	kɪp	mɪp	laɪn	slaɪn	dog
dont/his	ʃawt	tʃawt	rɪl	ɡɪl	dream
new	wʌrk	vʌrk	θɪn	plɪn	dress
their	fɔrθ	bɔrθ	hɛn	kɛn	drink
old	pɔt	drɔt	taɪl	flaɪl	dust
will	hɑrm	zɑrm	θɪrɪl	tʃɛl	film
want	strɛs	bɛs	stɔr	blɔr	food
did/my	ʃɛm	wɛm	pɔr	plɔr	friend
may	pæk	fræk	ʃɛr	plɛr	juice
they	bʌmp	ʃvmp	heŋ	meŋ	key
that/may	mɪn	flɪn	skɛr	grɛr	kid
will	kætʃ	brætʃ	skɔr	jɔr	kiss
have/can	nais	hais	str	vɪr	milk
one/will	klæs	slæs	blɛr	ɡlɛr	mind
tough	wʌlf	tʌlf	drɪl	rɪl	mouth
did	pɪk	fɪk	slɑr	skɑr	name
loud	dʒɪm	zɪm	flɔr	zɔr	noise

Appendix B (continued)

Word0	Critical Words				Target Word
	Word- Illegal	Nonword- Illegal	Word- Legal	Nonword- Legal	
that	kjut	grut	dɪm	ɡlɪm	page
much	lɛɡ	ɡlɛɡ	θvɪm	fʌm	pain
bright	brawn	shawɪn	stɑr	slɑr	pair
bad	dɾʌɡ	ʒʌɡ	ruwɪm	ʃuwm	pass
torn	stʌk	θʌk	ɡʌm	fʌm	piece
new	flæg	græg	tɔr	slɔr	plane
very	dɑrk	wɑrk	ɡrɪm	fɪm	play
chipped	dɛn	vɛn	ʃɛl	vɛl	plaque
will	tʃɛk	vɛk	sɛl	ɡɛl	pool
stained	rɛd	kɛd	dʒæm	zæm	plate
large	ræk	græk	mæs	tæs	price
feel	ɡrɛt	flet	hɒm	fɾɒm	pride
buy	klʌb	brʌb	blæk	dæk	salt
have/can	dɾʌm	sprʌm	rɪp	kɾɪp	seat
new/did	stɛm	grɛm	fɪp	fɪp	seed
new	tɪm	kɪm	læmp	flæmp	sale
the	bæg	glæg	hɒl	tʃɒl	tax
tough	mæp	væp	skɪl	θɪl	test
eat	lʌntʃ	ɡʌntʃ	ɡrɪl	dʒɒl	toast
hear	rɪtʃ	ʃɪtʃ	smɒl	fɾɒl	tone

APPENDIX C

STIMULI USED IN PART II (EXPERIMENT 6)

Word0	Critical Words				Target
	Word-Illegal	Nonword-Illegal ^{a,b}	Word-Legal	Nonword-Illegal ^{a,b}	
that	telescope	bɛləslop	bicycle	frobəmɪl	camp
nice	figurine	mɪpjərvin	cellular	ʃæljəlɹ	case
old	permanent	tʃɛrmɪgənt	courteous	bɛrvɪəs	cast
large	landscape	pændkrep	adventure	mʌgrɛntʃɹ	coast
high	envelope	gɛnbɪtrɒp	daffodil	fænjdʌrɪl	cost
huge/will	entertain	fɛnpɛrgɛn	religious	bətɪrəs	crowd
had/will	examine	mɪksæmɪn	personnel	vɛrgənɛl	cut
can	organize	bɔrdɛlaɪz	quarantine	glɔrɛndzɪn	dance
late	certificate	lɔrbɪsəkɪt	chaperone	tʃæpəron	date
his	terrific	pɛrlɪbɪk	assassin	lɔpævɪn	deal
wipe	geranium	bɛrmeziəm	dandelion	tʃæmpəfaɪən	dirt
this	handicap	fæmdəvɛp	vegetarian	tubətʃarɪən	dog
will/short	accurate	rægjəbɪt	imagine	ɪbævɪn	dream
new	exotic	lɛkmədʒɪk	feminine	lɛbəmɪn	dress
their	calcium	kælsiəm	tangerine	bæpməɪn	drink
old	cabinet	sækəbɪt	molecule	tələvjul	dust
sell	uniform	ʃuwbəwɛrm	controversial	fəŋgrəbɛrʃl	film
buy/will	sacrifice	bæklədais	labrador	zɪprəmɔr	food
my	mexican	lɛkrədʒɪn	gladiator	trædzɪɛtr	friend
that	energetic	bɛkɛrmɛdɪk	refrigerator	məgrɪdɛrɛbr	juice
give	limousine	ʃɪnəlin	versatile	ʃɪvməbəl	key
that	discipline	vɪʃəblɪn	muscular	rʌmyulɹ	kid
may/give	encourage	mɪntʊɪj	sinister	vɪnəskɪr	kiss
can/buy	fabulous	læbrələs	administer	bædlɪvɛstr	milk
one	devious	pɛgɪʃəs	predator	grubɪdʒɹ	mind
large	photograph	mɔnjutræf	animal	sæbəkul	mouth
has/did	democratic	dʒɛməbrætk	volunteer	bəmɪndɪr	name
loud	stadium	kleviəm	radiator	ʃevɪevɹ	noise
that	alphabet	mʌlʃəvɛt	curriculum	lubikeyədəm	page

Appendix C (continued)

Word0	Critical Words				Target
	Word-Illegal	Nonword-Illegal ^{a,b}	Word-Legal	Nonword-Illegal ^{a,b}	
feel	appendage	məpɛndɪdʒ	medium	segiəm	pain
bright	moccasin	fəgətɪn	designer	skəbaɪvɜ	pair
bad	synagogue	bɪnələg	auditorium	bɒdɪgɔriəm	pass
torn	acoustic	krəbustɪk	museum	sudiəm	piece
small	immaculate	pɪmædʒəlɪt	modular	pɑmyədɜ	plane
bad	academic	mæsəbɛmɪk	millenium	bərəgiəm	play
chipped	kindergarten	sɪmlɪrmərkɪn	commercial	səbɜrvɪl	plane
my	olympic	fəmlɪfɪk	crocodile	lɑbədzɑɪl	pool
stained	marinade	gerəbed	hologram	gɑvyəkrɛm	plate
large	almanac	bɒlnəmæk	thesaurus	pəʃbəs	page
feel	document	ʃɑkyələnt	coliseum	mɑləpiəm	pride
buy	abrasive	ləkrɛnɪv	arabic	θerəkɪk	salt
sell	aluminum	nələbɪnəm	scholarship	træbənɪp	seat
new	opium	ʃofiəm	cantaloupe	bæpərəp	seed
that	catalog	sænəmɔg	official	tʃəlɪʃəl	tax
may/tough	develop	rəmeləp	medical	revəmyul	test
make	expensive	mɪkfɛnsɪv	liberal	pɪmərɪ	toast
hear	average	tæpərɪdʒ	spiritual	trævəmyul	tone

^a Nonwords are shown in IPA format

^b Stress pattern for nonwords is the same as for words within each phonotactic condition