1984

The role of prosody in perception of lexical stress.

Cynthia M. Connine
University of Massachusetts Amherst

THE ROLE OF PROSODY IN PERCEPTION OF LEXICAL STRESS

A Masters Thesis Presented

by

Cynthia M. Connine

Submitted to the Graduate School of the
University of Massachusetts in partial fulfillment
of the requirements of the degree of
Master of Science

February 1984

Psychology

THE ROLE OF PROSODY IN PERCEPTION OF LEXICAL STRESS

A Thesis Presented

By

Cynthia M. Connine

Approved as to style and content by:

Charles Clifton, Jr., Chairperson of Committee

Lyn Frazier, Member

Alexander Pollatsek, Member

James M. Royer, Department Head

Psychology

ACKNOWLEDGEMENTS

I would like to thank the members of my committee, Chuck Clifton, Lyn Frazier and Sandy Pollatsek for their advice and support.

Each member contributed unique and much appreciated input on the long road to the final product. A special mention goes to Chuck Clifton. Chuck provided many gentle urgings, without which, this thesis may never have been completed. My heartfelt thanks go to him for his patience, encouragement and good humor.

A large part of the responsibility for the development of these experiments lies with Maria Slowiaczek. Maria's work inspired an interest in the mysteries of prosody early in my graduate career. She also fed me, nurtured me and stayed up many late nights with me at Bell Labs preparing experimental tapes. Maria is the spirit behind this thesis and I will always be grateful to her.

There are many people I have met in graduate school that have been a source of personal and professional support. Of these I want to particularly thank my dear friends Albrecht Infhoff and Susan Lima for many small and large kindnesses.

ABSTRACT

The role of prosody in perception of lexical stress

February 19, 1984

Cynthia M. Connine, B.S., Northeastern University

Directed by: Professor Charles Clifton, Jr.

Prosody refers to stress, pitch and rhythmic information in spoken language. One explicit theory of how prosodic information is used in natural language comprehension was proposed by Martin (1972). Martin suggests that the rhythmic continuation of an utterance is constrained by the preceding prosodic context. One study (Huss, 1979) has shown that lexical stress is perceived in concordance with this preceding prosody. The experiments reported here follow up the Huss experiment. Specifically, two kinds of prosodic information are identified as potentially influencing decisions in Huss' experiment: sentence stress and sentence rhythm. Sentence stress refers to the pattern of stressed and unstressed syllables whereas sentence rhythm refers to the relative timing of stresses. In experiment 1, naturally produced sentences were cross-spliced to obtain conditions in which sentence stress and sentence rhythm information conflicted. The results indicated that sentence stress was the major source of information used by subjects. However, acoustic analysis of the experimental sentences revealed no rhythmic timing changes. Several hypotheses are discussed as possible explanations for this lack of rhythmic anticipation.

A second experiment was done in which sentence rhythm was introduced in naturally produced sentences via linear predictive coding. Again, sentence stress was the major influence. The implications of these results are discussed in terms of a model in which prosodic information is used to constrain downstream rhythmic patterns. However, the prosodic information need not necessarily be physical timing changes. Rather, a more abstract representation of prosody may be influential. This abstract representation can be charecterized, minimally, as the lexical stress pattern of the item preceding the target word.

TABLE OF CONTENTS

List of Tables

CHAPTER I

INTRODUCTION

In order to fully understand a typical sentence, the listener must take the complex physical signal and transform it into a representation of meaning. In spoken language, there are many processes which must occur to achieve this goal. The speech stream must be segmented into words and lexical items must be identified. The lexical items must be grouped and related in phrases. The meaning of phrases must be determined and integrated with other phrases in the sentence. In a larger context, the meaning of the sentence must be integrated into the conversation as a whole.

A fundamental prerequisite to spoken language comprehension is the identification of sounds and words from a physical acoustic signal. This process occurs in a remarkably short period of time despite considerable variation in the physical signal. The acoustic information that indicates a segment or word may vary depending upon many different acoustic factors. For example, segmental information that specifies a consonant differs dramatically with the vowel that follows that consonant (Delattre, Liberman & Cooper, 1955). Specifically, the acoustic signal associated with the syllable /di/ begins with a rapidly rising formant transition. The signal that specifies the syllable /du/ begins with a rapidly falling formant transition. Yet, listeners perceive both syllables as beginning with the phoneme /d/.

Different speakers vary in the way they produce the same word due to differences in the shape of the vocal tract and styles of speech. Furthermore, one speaker differs from time to time in the production of a particular word. Even when a speaker attempts to repeat a word in the same way, control of the articulators is somewhat imprecise and a different acoustic pattern is produced each time. Identification of words from the speech stream is further complicated by the fact that a word boundary can not in all cases be determined by obvious discountinuities in the signal. A number of principles are involved during the production of fluent speech that preclude separation of words by pauses. For example, American English speakers palatize alveolar obstruents when followed by a palatal (eg. did you-- /diju/) (Oshika et al, 1975). When the same phoneme occurs at the juncture of two words, that pair of phonemes reduces to one phoneme (eg. real love-- /rielov/). In addition, there are instances when there is potentially more than one segmentation of any one utterance (eg. real love, reel of).

Contextual changes in the speech signal have posed considerable problems for language processing theories. In order to account for the seeming ease of comprehension, current processing theories have had to appeal to the active use of the listener's knowledge of the language to decode the signal. Through experience, the language user internalizes regularities that exist in the language. The assumption is that these internalized knowledge systems are used to structure what is assumed to be an impoverished acoustic signal.

Concerns about the way these knowledge sources are used to organize and represent the speech stream during processing have dominated much of psycholinguistic research. However, one source of information that potentially provides some of the organizational information necessary for language comprehension is prosody. For the present, prosody can loosely be defined as the melody of language. Recent theoretical proposals have emphasized the use of the prosodic structure in the signal to facilitate the comprehension process (eg. Slowiaczek, 1981). In the next section, the prosodic properties of spoken language are more precisely defined and evidence available that indicates the language processing system uses prosodic organization are considered. First, knowledge systems that could be incorporated into a processing theory are described. Then, claims about how language constraints are used in current processing theories are reviewed.

Linguistic Levels

There are a number of potentially useful knowledge sources available to the listener that could be used to constrain the organization of the speech stream. The constraints imposed by information in the sentence as it is being processed could help a listener interpret the signal despite the variability in the acoustic wave. The sentence is potentially constrained by the listeners' knowledge of the phonetic, lexical, syntactic and semantic properties of the language.

Words consist of a succession of sounds or phonetic segments. The sounds that can be combined together to form syllables and words are restricted and a language user knows which phonetic combinations are possible. For example, there is no single syllable word in the English language that contains the sound combination /zg/. When these two sounds do occur consecutively in an utterance the listeners knowledge of legal phonetic combinations could be used to define a word or syllable boundary (eg. please go--/plizgo/; nosegay--/nozge/).

Lexical information refers to knowledge about individual words. Each lexical item may be classified as a noun, verb, adjective. Prefixes or suffixes can be added to form a related word (eg. un+clear) or a word of a new lexical class (eg. clear+ly). The listener knows what words are in the language as well as the possible forms of a word. This knowledge can constrain what words could possibly be heard.

Syntax refers to the grouping of lexical items into larger constituents according to their grammatical relations. Users of a language know what possible syntactic configurations exist in their languge (eg. an adjective or group of adjectives must be followed by a noun). The listener could use this knowledge to constrain what words are heard and what lexical meaning is chosen. For example, consider the following sentence.

1. Sarah saw dust under the bed.

Although the phonemic string <u>saw</u> <u>dust</u> could be interpreted as a noun, such an interpretation would be ungrammatical in subject position. The presence of the noun <u>Sarah</u> necessitates that the string <u>saw</u> <u>dust</u> be parsed as a verb followed by an object. Furthermore, even though the lexical item <u>saw</u> can be used as a noun or verb, the verb usage must be assigned.

The study of the meanings of words and sentences is semantics. A language user knows the meanings of words and how to determine the meanings of the words. What is known to be true about the world can be used to determine the plausibility of an utterance. A speaker of English knows sentence 2 is not a plausible utterance.

2. The table ate the spaghetti.

A listener could use semantic information to set up expectations about what is likely to occúr later in a sentence.

Linguists recognize a fourth classification of knowledge in language description. This body of knowledge is termed pragmatics. Morris (cf. Lyons, 1977) defines pragmatics as dealing with "the origin, use and effects of signs within the behavior in which they occur". There are various formulations of the scope of pragmatics in linguistic theories. However, the basic notion is designed to focus on an utterance within a language-user context. This may include inferences such as that described in the following dialogue in a restaurant.

Customer: Do you have coffee to go?
Waitress: Cream and sugar?

In this discourse, the waitress assumes the customer's question was in fact a request for coffee to go without an explicit statement to that effect.

The listener's knowledge about the phonetic, lexical, syntactic, semantic and pragmatic properties of the language could be used during processing to constrain possible continuations of the utterance. Ambiguity due to imprecise articulation, noisy environments and contextual variations would be less disruptive if such knowledge was used since processing would be less dependent on an analysis of the acoustic waveform.

Language Processing Theories

Language processing theories have made different claims about how these language constraints are used. According to Marslen-Wilson (1978) sentence processing proceeds in a totally interactive manner. Constraints from all knowlege sources derived from prior items guide processing of the current item. The processing system uses syntactic and semantic information for lexical identification. Variability in the waveform is less disruptive because lexical identification depends on other sources of information to constrain possible lexical items.

In support of this model, Marslen-Wilson (1975) found that fast shadowers were more likely to restore a mispronounced lexical item when the syntactic and semantic environment was more highly constrained. These findings were interpreted as suggesting that information from all sources is available and influencing analysis of the current word processed.

An alternative view of the language processing system is that information is processed in a sequence of discrete stages of analysis (Forster, 1976). The process is hierarchically structured such that information that constitutes the output of a lexical processor feeds into a syntactic processor. Similarly, syntactic information is input to a message level processor. The output of an earlier stage can not be modified by a later stage of analysis. The way in which words are grouped into syntactic phrases is not modified by semantic properties of a sentence. Forster's model implies that each possible syntactic organization is computed, regardless of semantic plausibility. For example, one structure of sentence 3 would assign the phone as the direct object of the verb pacing.

3. WhileGeorge waspacing the phone rang.

This analysis would be considered even though this is a semantically implausible organization.

Similarly, the identification of words in Forster's model is not influenced by the way in which prior words in the string are grouped into phrases based on information concerning the meaning of the sentence. Lexical identification occurs as a discrete stage when sufficient phonemic information to specify a word is analyzed.

In order to account for contextual influences in sentence processing, Forster (1979) proposes a general problem solver (GPS). The GPS receives input from the lexical, syntactic and message processors, but cannot interfere directly with the operation of any of the processors. This system is the decision maker and unlike the processors from which it receives input, has access to general world knowledge and beliefs. The relative accessibility of information to the GPS, that is information transfer rates of the processors, determines the basis upon which the GPS makes its decision. The lexical information is analyzed and submitted to the syntactic processor prior to any analysis by the message processor. However, the transfer of information to the GPS from the message processor occurs at a faster rate than the lexical processor. Thus, decisions about the input can be based on message level information even though lexical information was actually processed first.

The language processing theories that have been described emphasize the use of internalized systems of linguistic knowledge during sentence processing. The listener imposes an organization of a sentence by appealing to his or her knowledge of regularities in the language. Lexical, syntactic and semantic knowledge is used in various ways to segment the signal into words, organize words into syntactic units and determine the meaning of an utterance. These processes involve active use of what the language user knows about possible organizations of sentences.

Prosody

It is possible that the prosodic aspects of an utterance may provide some of the organization necessary to interpret the speech stream. Prosody refers to intonation, rhythm and stress information in spoken language. Intonation refers to a gradual change in pitch over the course of a sentence or phrase. A falling intonation can be illustrated to the reader by saying the following sentence.

4. AlthoughLouisadidn't want to move to Indiana,she found she rather liked it.

When this sentence is produced by a speaker, a gradual drop in pitch can be noticed from the beginning of the sentence to the end of the first phrase (after Indiana).

Changes in the duration of syllables and placement of pauses are part of the overall sentence rhythm. For example, speakers typically lengthen the final syllable in a phrase (Klatt & Cooper, 19758; Cooper, Paccia & Lapointe, 1978). Phrase final lengthening and pauses are part of the overall sentence rhythm.

Bolinger (1964) introduced a distinction between accent and stress. Accent refers to the actual occurrence of a pitch movement such that this portion is significantly different from the surrounding pitch. For example, in sentence 5, shoe receives nuclear accent (if produced as in response to the question "what kind of boxes does he make?").

5. He's a shoe box manufacturer.

Stress is considered to be an abstract property of lexical items. Two forms of the same phonemic string can be distinguished by word stress. Consider for example, the pair convert vs. convert. Primary stress is indicated by the underlined segment. In production, the noun form of the word is signaled by the presence of primary accent on the first syllable as in the following example:

6. Jules was a convert to the Reunification church.

Primary accent on the second syllable indicates the verb form:

7. Pam wanted to convert their heating system to a wood burning stove.

In general, stress is the potential for a syllable of a word to receive accent. The lexical item cognac has stress on the first syllable since it is this syllable that may potentially receive a pitch movement in production. In sentence 8 (produced in response to the question "does your sister drink cognac after dinner?"), the lexical item vermouth receives nuclear accent on the second syllable. The stressed syllable of a lexical item may receive a pitch movement in other than nuclear accent position. This type of pitch movement has been called a pitch accent (Pierrehumbert, 1980). Thus in sentence 8, the first syllable of sister receives a pitch accent.

8. No, my sister drinks vermouth after dinner.

Thompson (1980) defines a third prosodic property of syllables, that is salience. Salience of a syllable is dependent upon the rhythmic unit, the foot. A foot begins with a salient syllable, that is, an acoustic realization of some complex pattern of increased duration, intensity and fundamental frequency. However, Thompson also hypothesizes that perception of a salient syllable is not always dependent upon explicit acoustic cues. Perception of the rhythmic unit of the foot determines which syllable is perceived as salient. Thus, a salient syllable may be perceived in the absence of any physical marker in the signal.

Stress is typically signaled in the acoustic waveform by a complex pattern of information that includes duration, intensity and fundamental frequency (cf. Gay,1978). Although each of these individual cues is a sufficient cue for stress, it is unclear which cue is most important. Nakatani and Aston (1978) reported that duration is a better perceptual cue to primary lexical stress while others have found pitch is the most important cue (Fry, 1958).

Research on the role of prosody during language comprehension has been aimed at a number of levels of procesing the signal. Some lines of research have focused on determining the role of prosody in segmentation of the speech stream into words and syntactic units. Other studies have tried to demonstrate that prosody serves to maintain intelligibility and is a source of continuity in the speech signal. Performance on phoneme monitoring tasks has been used to provide evidence that prosody highlights important points in the signal in order to direct attentional capacity. One specific proposal has claimed that the patterning of stressed segments constrains the timing of the continuation of an utterance (Martin, 1972). Each of these areas of research suggest that prosodic information in the signal provides structure that is used by the language processing system and will be considered in turn.

Classification of sounds that differ in one distinctive feature have been found to be influenced by the surrounding rhythmic environment. Port (1978) found that perception of the voicing dimension of the medial stop consonant in the word pair rapid vs. rabid (/p/ vs. /b/) changes systematically with the rate of the carrier sentence. Specifically, as a speaker's articulation rate (average speaking rate per syllable) decreases, more silence is needed to cue the voiceless consonant /p/. Miller & Grosjean (1980) replicated these findings and showed in addition that articulation rate is weighted more heavily than pause structure (average amount of pause time in the sentence per syllable) in influencing perception of this segmental distinction. They conclude that sentence rate and to a lesser degree the pause structure in their sentences was used by listeners when making segmental judgements.

Recent work has shown that prosody provides information about the segmentation of an utterance into words. Nakatani & Shaffer (1978) had speakers produce sentences in which tri-syllabic phrases were mimicked by nonsense syllables. For example, in sentence 9 the phrase <u>remote stream</u> was produced as /mama ma/.

9. The remote stream was perfect for fishing.

When these reiterant phrases were presented in isolation, listeners could correctly parse the phrases into the intended word configuration.

Other studies have shown that prosody can indicate the intended syntactic analysis in sentences that have more than one possible structure. Lehiste (1973) instructed speakers to produce sentences such as 10 with the intention of communicating one of the two possible meanings.

10. The old men and women stayed home.

Speakers differentiated the intended segmentation of these sentences by inserting pauses, increasing duration and laryngealization at the intended phrase boundary. Listeners could reliably choose the intended analysis of these sentences. When duration was systematically manipulated by computer, listeners could still reliably choose the intended meaning (Lehiste, Olive and Streeter,1976#). It should be noted that not all ambiguities could be differentially produced. Sentences with both deep structure and surface structure ambiguities could not be reliably disambiguated.

It has also been reported that intelligibility of speech is enhanced when prosodic information is present. Huggins (1978) compared the intelligibility of synthesized sentences in which either timing, pitch or both were distorted. He found that sentences with abnormal pitch were substantially less intelligible than those in which pitch and timing were maintained (86% vs. 63% words correct). When only timing information was disrupted there was a further decrement in words reported correct (52%). Distortion of both kinds of information decreased performance only slightly more (48% words correct). Dooling (1974) found that prior exposure to a particular sentence rhythm decreased perception of words in a sentence presented in noise if an opposing rhythmic set was induced. For example, a group of sentences were presented with the prosodic pattern of sentence 11. In sentence 11, the prosodic pattern can be charecterized as an alternating series of strong and weak syllables (s=strong, w= weak).

11. They are sneaky foxes.
    s w s w s w

When the final sentence in the group differed in prosodic pattern (eg. sentence 12) perception of this sentence was disrupted.

12. They are severe defeats.
    w s w s w s

Dooling suggested that a rhythmic set was induced by the prosodic pattern of the initial group of sentences. This rhythmic set interfered with perception of the final sentence that did not have a parallel rhythm to this set.

Darwin (1975) has suggested that prosody provides a continuity in speech and guides the listeners attention to the ongoing information flow in the acoustic signal. He found that subjects performing a dichotic shadowing task had intrusions from the unattended channel when the prosody was switched to the unattended channel.

There is evidence that suggests that performance in some tasks is facilitated at points in the speech signal that are stressed. Phoneme reaction time is faster to targets in stressed items than unstressed items (Cutler & Foss, 1977; Shields, McHugh & Martin, 1974). Moreover, this effect was not simply due to a strong or more clearly articulated signal in these positions. When a neutrally stressed item was inserted into sentence stress position, phoneme reaction time to these targets was still faster (Cutler, 1976; see also Cutler & Fodor, 1979[1]).

These studies suggest that listeners allocate increased attentional capacity in anticipation of a stressed syllable. One way that a stressed syllable could be predicted from the prior prosodic context is if the speech signal is isochronous, that is, if the onsets of stressed syllables are approximately equally spaced in time. However, Lehiste (1973) has reported that speech is not strictly isochronous. She measured interstress intervals in recorded sentences according to a set of criteria proposed by Abercrombie. This system was used to determine the duration of ischronous rhythmic units or rhythmic feet. She found clear differences in the duration of rhythmic feet. Furthermore, listeners had great difficulty in judging these differences in duration in a sentence context. If the corresponding durations were presented as noise bursts, performance improved: listeners were better at judging differences in duration when performing this task. Lehiste suggests that listeners impose isochrony on the speech signal.

Donovan & Darwin (1979) have reported similar findings. Listeners were asked to match the rhythm of a series of noise bursts with the rhythm of sentences. It was found that the durations between the noise bursts manipulated by listeners were more regular than speech. These deviations from veridical perception were not found when the sentence contained an intonational boundary.

Although listener's judgements of duration in these tasks tend to deviate from the actual speech toward isochrony, people can anticipate with some precision the occurence of a stressed syllable. Martin (1979) recorded six nonsense syllables in an alternating stressed-unstressed pattern. The vowel duration of one of the items was either lengthened or shortened. The consequence of this manipulation was to disrupt the rhythm of the sentence. He found that phoneme monitoring reaction time to a target as far as four syllables downstream from a distorted item was significantly slowed. These data indicate that listeners are sensitive to small changes in timing. Furthermore, these effects ocurred on non-adjacent items indicating a dependency between these segments. Martin suggests that a rhythmic expectancy component operates to constrain the prosodic continuation of an utterance. Presumably, processing is disrupted when expectancies resulting from prosodic constraints conflict with the information currently available in the signal.

CHAPTER II

EXPERIMENT 1

One study has reported that identification of words with ambiguous lexical stress patterns is influenced by the surrounding rhythmic environment (Huss, 1979). In this experiment, two carrier sentences were recorded with nuclear accent on the underlined item.

13. WhatdoesEnglish(pendant) mean.

14. Say the word (defeat) again.

According to Huss, the alternating sequence of stressed and unstressed syllables of sentence 13 predicts a lexical item with primary stress on the first syllable in the position indicated by the parenthesis. Similarly, the pattern of sentence 14 predicts stress on the second syllable in the target position.

To test this, the target words import and decrease were inserted into the target postion in both sentences (eg.pendant was removed and the appropriate target word was inserted). There were two versions of each word: one version was taken from a sentence context where the target word had been used as a noun, while the second version was excised from a sentence frame where the word had been used as a verb. The sentences from which the target words were taken were recorded such that the target words occured immediately after the nuclear accent position (post nuclear position). The effect of this manipulation is that target words were ambiguous in perceived lexical stress.

The mosaic sentences were presented to listeners and their task was to indicate the lexical stress pattern of the target item. Huss found that listener's judgments of lexical stress were influenced by the overall rhythmic pattern of the carrier sentences. Both versions of the words _import_ and _decrease_ tended to be perceived as having lexical stress on the first syllable in sentence 13 (approximately 61%). In a similar fashion, approximately 59% of the target words inserted into sentence 14 were perceived with primary lexical stress on the second syllable. That is, listeners tended to perceive lexical stress in a way that maintained an alternating stress utterance. Target word judgements conformed with a general tendency to perceive stressed segment after an unstressed segment.

However, listeners' judgements of lexical stress may have been influenced by two kinds of rhythmic information in the sentences in this experiment. The first kind of information concerns the pattern of strong/weak syllables. I will call this the _sentence stress pattern_. The carrier sentences always conformed to an alternating pattern--the word _pendant_ did not deviate from the weak/strong pattern of the sentence it was recorded in. In a similar fashion, the word defeat conformed with a stong/weak sentence stress pattern.

The second kind of information concerns the timing of segments in the carrier sentences. Consider the following sentence.

15. Say the word pendant again.

In this sentence, two strong stresses occur on adjacent syllables. The sentence stress pattern violates a strong/weak alternation. Rhythmic changes in anticipation of the weak-strong-strong stress pattern may occur. This will be referred to as sentence rhythm.

In fact, Huggins (1978) reported that changes in duration due to sentence stress pattern can span word boundaries. For example, a stressed syllable may be shortened when followed by an unstressed syllable. The duration of shout in the following sentence

16. We shout aloud.

is shorter than in

17. We shout loud.

In the Huss experiment both sentence stress and sentence rhythm may have acted to influence listeners judgements. In the first experiment I will replicate the Huss experiment. In addition, I will attempt to separate the effects of sentence stress and sentence rhythm. For example, in sentence 15, if listeners' perception is influenced only by the sentence stress pattern, a word inserted into the target position would be perceived with stress on the second syllable. However, if sentence rhythm influences perception, then a target word would tend to be perceived with stress on the first syllable. Acoustic analysis of the carrier sentences will be done to determine differences in production.

Another concern is the ambiguous nature of the target words. The presence of two adjacent stressed syllables may also alter the production of the target words. In Huss's experiment, target words were recorded in post nuclear position and in sentence contexts with alternating stress to minimize the acoustic differences between noun and verb forms. In this experiment target words will be produced in environments similar to the Huss experiment. In addition, sentences will be constructed in which the target word interrupts an alternating pattern. These target words may be less ambiguous than those produced in sentences with alternating stress, that is these words may have lexical stress emphasis. If so, prosodic information in the carrier sentences may be less influential in listener's judgements of these words.

Four target words were chosen. Two of these words (CONVICT and CONDUCT) involved relatively more vowel reduction when stress is shifted from first to second syllable. These words undergo a greater degree of vowel change than the remaining target words (DECREASE and DISCOUNT) and thus may provide more segmental information as to the stress pattern. This charecterization of the target words is supported by perceptual evidence. Taft (1980) presented listeners with the initial syllable of bisyllabic words and asked them to indícate lexical stress. She found that listeners could identify stress more reliably if the stress shift involved vowel reduction. Furthermore, Cutler and Clifton (1982) found that performance on a phrase grammaticality judgement task was disrupted to a greater extent when listenees heard <u>to CONvict</u> versus <u>to PERmit</u>. These experiments indicate that vowel reduction in words like CONVICT and CONDUCT provides additional cues to lexical stress. Thus, perception of thesewords may be influenced to a lesser extent by contextual rhythmic factors.

To obtain an indication of the ambiguity of the target words when embedded in sentences, listeners will be asked to rate each word for salience of stress information.

Method

<u>Subjects</u> Forty four students at the University of Massachusetts served as subjects.

Materials and design Two sentence frames were constructed and are listed in table 1. The target word position is indicated by the underlined portion. Sentence 1 was constructed to have an alternating weak-strong sentence stress pattern. This pattern is indicated below sentence 1 by the symbols *w* (weak stress) and *s* (strong stress). Sentence 2 was constructed to form an alternating strong-weak pattern, as indicated in table 1.

Four word pairs were used. These are listed in table 2. Two of these items receive relatively little vowel reduction when produced with accent on the second syllable (DECREASE, DISCOUNT). The remaining two words receive relatively more vowel reduction when accent is on the second syllable (CONVICT, CONDUCT). Each word was recorded in each sentence frame resulting in two original recordings for each word. Thus a total of 16 naturally spoken sentences were recorded. In all sentences, nuclear accent stress was placed on the word immediately prior to the target word.

Table 3 shows stress patterns for each of the four original recordings. For ease of explication, the word *discount* will be referred to in the following section. In

Table 1

1.I think if you recite * to me then I will know the word.
w s w s w s w s w s w s w s

2.When youtryto parrot * quickly I will sayit slowly.
s w s w s w s w s w s w s w

Table 1. Sentence frames used in Experiments 1 and 2.

Table 2

CONvict	conVICT
CONduct	conDUCT
DEcrease	deCREASE
DIScount	disCOUNT

Table 2. Target words used in Experiments 1 and 2.

sentence 1 and 3, the sentence stress pattern has an alternating weak-strong pattern. If perception of the target word is influenced by the alternating pattern, then in sentence 1 the target word would tend to be perceived with stress on the first syllable. Sentences 2 and 4 have a non-alternating sentence stress pattern. However, since these sentences are naturally spoken, production of these sentences anticipates the non-alternating pattern. The listener may be able to use this rhythmic pattern to identify the stress of the target word.

These sentences were cross-spliced to create 12 additional conditions. The 16 naturally spoken recordings were digitized, segmented, cross-spliced and resynthesized using the Bell Laboratories PCM system. Each original recording was divided into three segments. These three segments are illustrated in table 4. The target word segment was divided immediately before the release of the stop consonant (i.e. before /d/ in discount). The final division was made immediately before the release of the stop consonant of the word following the target word (i.e. before the /t/ in the word to). The segments were recombined to form the original natural sentences or sentences with the target word segment from one of the alternate versions. The complete set of cross spliced

Table 3

1.I think if you recite disCOUNT to me

w s w s w s w s w s

then I will know theword.

w s w s w s

2.I think if you recite DIScount to me

w s w s w s s w w s

then I will know the word.

w s w s w s

3.When you try to parrot DIScount quickly

s w s w s w s w s w

I will say it slowly.

s w s w s w

4.When you try to parrot disCOUNT quickly

s w s w s w w s s w

I will say it slowly.

s w s w s w

Table 3. Original sentence productions used in experiment 1.

Table 4

1f 1t 1f

1.I think if you recite /disCOUNT/ to me then...

2f 2t 2f

2.I think if you recite /DIScount/ to me then...

3f 3t 3f

3.When you try to parrot /DIScount/ quickly I...

4f 4t 4f

4.When you try to parrot /disCOUNT/ quickly I...

Table 4. Recombined segments used in experiment 1.

sentences resulted in 16 conditions which are illustrated in table 5. Four of these conditions were created by cross splicing the two _ws_ and the two _sw_ sentence frames. These conditions have _incompatible sentence rhythm_ because the sentence frame was produced to anticipate a lexical stress pattern that does not match the inserted target word.

Half of the target words were produced in non-alternating contexts which may have created _lexical stress emphasis_ on the target word to compensate for the non-alternating pattern. The remaining eight conditions were constructed by presenting each of the test sentences in two forms, with and without lexical stress emphasis on the target word. Eight of the total 16 conditions had _alternating sentence stress_ since the target word matched the _sw_ or _ws_ pattern of the frame. Similarly, the remaining eight conditions formed non-alternating sentence stress since the target word did not match the _sw_ or _ws_ pattern of the frame.

Two experimental tapes were recorded. Each tape contained a different randomization of the 16 conditions per word (a total of 64 experimental sentences) and 64 filler sentences. One half of the subjects listened to one tape. The remaining subjects heard the second tape.

Table 5

Combined segments	Sentence rhythm	Sentence stress	Lexical emphasis
1f 1t 1f	COMP	ALT	NE
2f 2t 2f	COMP	NON-ALT	E
3f 3t 3f	COMP	ALT	NE
4f 4t 4f	COMP	NON-ALT	E
1f 2t 1t	INC	NON-ALT	E
2f 1t 2f	INC	ALT	NE
3f 4t 3f	INC	NON-ALT	E
4f 3t 4f	INC	ALT	NE
1f 3t 1f	INC	NON-ALT	NE
2f 3t 2f	COMP	NON-ALT	NE
3f 1t 3f	INC	NON-ALT	NE
4f 1t 4f	COMP	NON-ALT	NE
1f 4t 1f	COMP	ALT	E
2f 4t 2f	INC	ALT	E
3f 2t 3f	COMP	ALT	E
4f 2t 4f	INC	ALT	E

Table 5. Conditions used in experiment 1.
f=frame
t=target
COMP=compatable, INC=incompatable sentence rhythm
ALT=alternating, NON-ALT=non-alternating sentence stress
E= lexical emphasis, LE=no lexical emphasis

<u>Procedure</u> The experimental sentences were played over Layfayette SP-55 stereo headphones at a comfortable listening level. Subjects were asked to listen to each sentence and indicate lexical stress of the target item. The target lexical item was shown on an answer sheet provided for each subject. Subjects indicated perceived lexical stress by underlining the appropriate syllable of the item on the answer sheet. A space for a rating response was also shown for each target. A two second interval of silence after each sentence provided time for subjects to make a response.

Results

An initial analysis of the proportion of responses for first syllable stress revealed a significant main effect of the four words and some complex interactions. Thus a separate analysis was performed for each individual word. Summary data for main effects for individual words are shown on Appendix A, table 13. A summary table for two way interactions is also presented (see Appendix A, table 14). Appendix D, table 22 presents results from the 16 original conditions.

Decrease

<u>Main effects</u> A repeated measures ANOVA showed a reliably greater preference for first syllable stress when sentence stress favored a first syllable target, relative

to second syllable sentence stress ($F(1,42) = 8.82$, $p <$ 01). The proportions of first syllable stress reports for a sentence stress favoring a first syllable and a second syllable target was 43% and 33%, respectively. There was no significant effect of sentence rhythm ($F(1,42) <1$). The overall effect of intended word stress was significant ($F(1,42) = 30.12$, $p < .001$). The first syllable target was perceived as stressed on the first syllable more often than a second syllable target (48% and 28%, respectively).

Interactions A target x sentence stress interaction ($F(1,42) = 8.86$, $p < .01$) revealed that sentence stress influenced reports of first syllable targets to a much greater extent than second syllable targets. These values are shown in Appendix 1, table 14.

A significant target x exaggeration interaction ($F(1,42) = 17.92$, $p < .0001$) indicated a larger proportion of first syllable stress reports for a first syllable non-exaggerated target. This pattern did not hold for the second syllable target (33% and 2% difference between exaggerated and non exaggerated versions, respectively).

Inspection of a significant sentence stress x exaggeration x target interaction ($F(1,42) = 9.93$, $p <$ .01) revealed the following pattern of data (see table

Table 6

	Target			
	First		Second	
	Exaggeration			
	Y	N	Y	N
Sentence Stress				
first	37	78	31	26
second	26	51	22	32

Table 6. Percentage of first syllable responses for lexical item Decrease (target x exaggeration x sentence stress three way interaction). Experiment 1.

6). Sentence stress was more influential if a first syllable target was not exaggerated (27%) than if it was exaggerated (11%). An exaggerated second syllable target was slightly influenced by sentence stress (9%). However, there was a greater number of first syllable reports of the non-exaggerated second syllable target if sentence stress was compatible with a second syllable target (6%). It is possible that cues to words stress were neutralized in the non-exaggerated version to a greater extent for first syllable productions than second syllable productions. Conversely, second syllable productions may receive more pronounced exaggeration than first syllable productions.

Additional significant interactions included an order x target interaction ($F(1,42) = 4.16$, $p < .04$) and a three way order x sentence stress x sentence rhythm interaction

($F(1,42) = 7.84$, $p < .01$).

Discount

Main Effects An identical ANOVA for the lexical item discount showed similar results. Sentence stress reliable influenced reports of word stress ($F(1,42) = 13.9$, $p < .001$). A greater proportion of first syllable stress reports was found for a second syllable sentence stress than a sentence stress favoring a second syllable stress (57% vs. 42%, respectively). The effect of sentence rhythm was marginally significant (F(1,42) = 2.27, p < .10). There was a greater proportion of first syllable reports when sentence rhythm favored first syllable stress relative to a second syllable sentence rhythm (52% vs. 47%, respectively). Overall, a greater proportion of first syllable productions were reported as such than second syllable productions (59% vs. 40%, respectively, F(1,42) = 19.67, p < .0001).

Interactions The target x exaggeration interaction showed a somewhat different pattern for discount than decrease (F(1,42) = 7.32, p < .01, see Appendix A,table 14). Non-exaggerated first syllable targets showed proportionally greater first syllable reports than the exaggerated versions (63% vs. 55%, respectively). However, exaggeration tended to reduce the target effect for second syllable targets (35% vs. 44%, non-exaggerated and exaggerated, respectively).

Additional significant effects included a three way interaction of order x sentence stress x sentence rhythm (F(1,42) = 4.99, p < 03). Two four way interactions were found: order x sentence stress x sentence rhythm x target (F(1,42) = 8.61, p < .01) and sentence stress x sentence rhythm x target x exaggeration (F(1,42) = 7.2, p < .01).

Convict

Main effects Similar results were also found for the lexical item convict (see Appendix A, table 13). A sentence stress that favored a first syllable stress resulted in more first syllable target word reports than a second syllable sentence stress (46% vs. 52%, respectively); $F(1,44) = 7.5$, $p < .01$). No main effect of sentence rhythm was found ($F(1,42) < 1$). Overall, a word produced with intended stress on the first syllable was perceived as such to a greater extent than a second syllable production (88% vs. 11%, respectively; $F(1,42) = 308.79$, $p < .01$).

Interactions A significant sentence stress x target interaction ($F(1,42) = 6.04$, $p < .01$) revealed the following pattern of data (see table 11). An intended first syllable target stress was reported as such for a sentence stress that favored both a first and second syllable target (88% vs. 89%, respectively). However, reports of an intended second syllable target were influenced by the surrounding sentence stress: a sentence stress that favored a first syllable stress target showed more first syllable reports that a sentence stress that favored a second syllable target (17% vs. 5 %, respectively).

The significant sentence rhythm x target x exaggeration interaction is shown in table 7 ($F(1,42) = 4.07$, $p < .05$). Exaggerated second syllable targets that appeared in a rhythmically compatible sentence were reported as such to a greater extent than in an incompatible sentence rhythm (9% vs. 17%). This effect was reversed for first syllable exaggerated targets: more first syllable stress words were reported in a second syllable sentence rhythm than a first syllable sentence rhythm (92% & 86%, respectively). The opposite pattern of data held for non-exaggerated targets.

Additional significant interactions included a three way sentence rhythm x sentence stress x exaggeration interaction ($F(1,42) = 4.63$, $p < .03$) and a four way order x sentence rhythm x target x exaggeration interaction ($F(1,42) = 4.07$, $p < .05$).

Table 7

	Exaggeration YES, Target 1	Exaggeration YES, Target 2	Exaggeration NO, Target 1	Exaggeration NO, Target 2
Sentence Rhythm				
first	86	17	89	6
second	92	9	85	11

Table 7. Percentage of first syllable responses for lexical item Convict (Target x exaggeration x sentence rhythm interaction). Experiment 1.

Conduct

Finally, an ANOVA was done on the lexical item conduct. A design error was discovered in the first experimental order. Thus, only the data from the 22 subjects tested on the second order of experimental sentences were considered.

Main Effects A significant main effect (see Appendix A, table 13) for intended stress was found ($F(1,21) = 202.07$, $p < .001$). An intended first syllable stress target was reported as such more frequently than a second syllable target (91% vs. 9%, respectively). A significant main effect of exaggeration ($F(1,21) = 6.6$, $p < .01$) revealed that non-exaggerated targets were reported as first syllable stress more frequently than exaggerated targets (2% vs. 47%, respectively). No main effect of sentence stress ($F(1,21)$ $p < 1$) or sentence rhythm ($F(1,21) < 1$) was found.

Rating Responses

After subjects labeled each stimulus as first or second syllable stress, each stimulus was rated on a scale 1 to 6. The number 1 was assigned to stimuli with clear, unambiguous first syllable stress and a 6 to unambiguous second syllable stress words. The number 2 or 5 was assigned to those stimuli that were unambiguously perceived as first syllable or second syllable stress, respectively, but these categories were for stimuli that were not as clearly stressed. Finally, the number 3 or 4 was assigned to those stimuli that were perceived as ambiguous first or second syllable stress, respectively. The ANOVA's reported are in every respect identical to the previous analyses. A separate analysis was performed for each individual word. Summary tables 15 and 16 in Appendix A present main effects and two way interactions, respectively. In addition, appendix D, table 23 presents the data from the original 16 conditions.

Decrease

Main Effects A repeated measures ANOVA for the target word decrease revealed three main effects: sentence stress ($F(1,42) = 6.25$, $p < .01$), exaggeration ($F(1,42) = 43.98$, $p < .0001$) and target ($F(1,42) = 30.34$, $p < .001$). The target effect indicates that although, in gerneral, subjects assigned stimuli to the appropriate categories, these stimuli were perceived as relatively ambiguous. A first syllable target was rated as an ambiguous first syllable stress (3.4). Similarly, a second syllable target was rated as an ambiguous second syllable stress (4.27).

Interactions Although the proportion data did not show a significant target x sentence rhythm interaction, the rating responses proved to be significant ($F(1,42) = 3.99$, $p < .05$). Subjects gave a slightly more confident rating to a first syllable target if the sentence rhythm was compatible with the target: a first syllable target was rated slightly less ambiguous if it appeared in a sentence rhythm that was compatible with a first syllable stress, relative to a second syllable sentence rhythm (3.44 vs. 3.54, respectively). This pattern of data did not hold for a second syllable target: a compatible sentence rhythm yielded slightly less confident ratings than an incompatible sentence rhythm (4.17 vs. 4.37, respectively).

Similar to the proportion data, two additional two way interactions were found: target x exaggeration ($F(1,42) = 18.29$, $p < .001$) and sentence stress x target ($F(1,42) = 7.84$, $p < .01$). Finally, two three way interactions were found significant: order x target x sentence rhythm ($F(1,42) = 12.27$, $p < .001$) and target x exaggeration x sentence stress ($F(1,42) = 9.43$, $p < .01$).

Discount

Main Effects As was the case for the proportion data, the lexical item discount showed a main effect of sentence stress ($F(1,42) = 15.69$, $p < .001$) and target ($F(1,42) = 26.41$, $p < .001$). The target effect indicates that similar to decrease, this lexical item was perceived as relatively ambiguous (3.1 and 3.9, first and second syllable stress respectively).

Interactions A sentence rhythm x target interaction (F(,42) = 5.84, p < .02) revealed that a first syllable target was rated more confidently as such if it appeared in a first syllable sentence rhythm, relative to a second syllable sentence rhythm (2.9 vs. 3.2, respectively). A second syllable target was rated slightly more confidently as such if it appeared in a sentence rhythm favoring a first syllable stress, relative to second syllable sentence rhythm (4.0 vs. 3.82, respectively). This interaction was not significant in the proportion data. Additional significant effects included a two way target x exaggeration interaction (F(1,42) = 5.96, p < .02), a three way sentence rhythm x sentence stress x target interaction (F(1,42) = 4.74, p < .03) and two four way interactions: order x sentence stress x sentence rhythm x target (F(1,42) = 5.,53, p < .02) and sentence rhythm x sentence stress x target x exaggeration (F(1,42) = 4.34, p < 04).

Convict

The analysis of the target word convict revealed a single main effect and two complex interactions. The significant main effect of target (F(1,42) = 190.4, p < .001) showed a first syllable stress target was rated unambiguously as a first syllable stress (1.89). Similarly, a second syllable target was rated quite confidently as such (5.07). Other significant effects included a three way sentence rhythm x target x exaggeration interaction ($F(1,42) = 8.94$, $p < .01$) and a four way order x sentence rhythm x target x exaggeration interaction ($F(1,42) = 7.39$, $p < .01$).

Conduct

Main effects The analysis of the lexical item conduct showed that a first syllable production was rated unambiguously as such (1.86) and a second syllable production was rated confidently as second syllable stress (5.14), ($F(1,42) = 128.57$, $p < .001$). Further, a significant main effect of exaggeration was found ($F(1,42) = 10.43$, $p < .01$) (see table 14).

Acoustic Analysis

Acoustic analyses of the experimental sentences were performed in order to determine if the hypothesized anticipatory changes did occur. A number of analyses were done and a representative example of the types of measurements done are shown on appendix B. Appendix B presents the total duration of the precursor context word and the lexical item in the original production (i.e. reCITE deCREASE). In addition, the duration of the context word is shown. In order to determine if any rhythmic changes did occur, the proprtional duration of the context word relative to the total duration of the phrase was computed. We expected that a ws context word produced prior to a sw lexical target (i.e. reCITE DEcrease) would be proportionally longer than when produced prior to a ws target (i.e. reCITE deCREASE). As can be seen from appendix 3, there is no evidence for this type of duration change in any of the lexical items. Other measures produced comparable results.

Discussion

With the exception of one lexical item (CONDUCT), Experiment 1 indicated that sentence stress reliably influenced listeners judgements of lexical stress. However, the hypothesized influence of sentence rhythm did not reach acceptable levels of statistical significance. Only in the case of the target word CONVICT did the effect of sentence rhythm approach marginal significance in the predicted direction. Thus, we can conclude that sentence stress was the dominant factor in the Huss experiment.

This conclusion must be qualified however, by a potential shortcoming of the stimuli used in Experiment 1. The presence of rhythmic changes in the experimental sentences is essential for the influence of sentence rhythm, independent from sentence stress. As previously mentioned, acoustic analysis of the experimental sentences (see Appendix B) did not reveal any rhythm related duration changes. Thus, we cannot conclude that anticipatory rhythmic information cannot be used since such information may not have been available.

Assuming that anticipatory rhythmic information does exist in natural speech (cf. Huggins, 1972), there are at least two explanations for its nonoccurance in these materials. One possible explanation may lie in the peculiar construction of the sentence frames. Specifically, the rather unnatural alternating pattern of stressed and unstressed syllables in the sentence frames may have reduced anticipatory rhythmic changes. Durational changes which anticipate the following stress pattern in natural speech may have been reduced or neutralized by the presence of such an extreme alternating pattern.

An alternative explanation for the lack of anticipatory timing changes may lie in the fact that the sentences were produced with nuclear stress on the stressed syllable of the precursor word. The sentences were produced in this way in accord with the design of the Huss experiment. It may be that anticipatory timing changes are reduced when a word is in nuclear position.

The available acoustic data do not allow us to determine which explanation, if either , is the most appropriate account for lack of anticipatory timing changes. It is possible, however, to create the desired timing changes in the experimental sentences by using linear predictive coding and resynthesized speech. This method would allow us to maintain sentence stress as a variable and obtain precise control of sentence rhythm.

CHAPTER III

EXPERIMENT II

A second experiment was conducted in which sentence rhythm was manipulated by varying the duration of the precursor word (reCITE, PARrot). The stressed syllable was chosen as the locus of the duration changes. This was based on the assumption that duration between stressed syllables is the critical rhythmic unit in this task (cf. Lehiste, 1973; Fowler, 1979).

The Bell Labs LPC system was used to create sentences that were compatible and incompatible with the stress pattern of the following word. A single recording of each sentence frame (see Table 1) was used as the base frame. In order to create a sentence frame that has compatible rhythm with a <u>ws</u> target word the duration of the stressed syllable of the precursor context word was shortened. Similarly, another version was created to form a compatible sentence rhythm with a <u>sw</u> word. Here, the duration of the precursor context word was lengthened relative to the base frame. The direction of the duration change is based upon empirical acoustic analyses performed by Huggins (1972). As previously mentioned, Huggins found that word duration is conditioned by the stress of the previous word. Specifically, word duration is lengthened if that word is preceded by a stressed syllable. The conditions created with these sentence frames constituted a partial replication of the first experiment. In addition, the base frame served as a condition against which the duration manipulation may be assessed.

The data from the first experiment also suggest that the degree of ambiguity of target word stress was a factor in determining the relative influence of sentence stress. The effect of sentence stress was reduced or absent for those words that were rated as relatively unambiguous. As a class, these unambiguous target words differed in terms of vowel reduction (conduct, convict). In the present experiment, we attempted to manipulate word stress by varying duration. Three versions of each target word were prepared. The durations of the stressed segments for the first version were altered to obtain unambiguous first syllable stress. A second version of each target word was prepared with unambiguous stress on the second syllable. Finally, a third version was ambiguous in terms of its stress pattern. Thus, influence of target word ambiguity may be assessed relative to the influence of the sentence frame. In addition, fundamental frequency of the test sentences was set at a constant value in an attempt to isolate the influence of timing information independent from changes in pitch.

Method

Subjects Forty four students at the University of Massachusetts served as subjects.

Materials and design One production of each sentence frame (ws and sw alternation, see table 8) was chosen from 8 sample recordings of each frame. The sample recordings (see Table 3). In order to obtain sentence portions prior and subsequent to the target word that were in the midrange for each test frame, the following measurements were done. The total duration of the utterance prior to the target word onset was measured. A production was chosen from the midrange of the sample for each frame. In addition, the total duration of the context following the target word offset was determined. Again, a midrange duration was selected as the base from each set of 8 recordings. The base durations are listed in table 6. Next, the duration of the precursor context word (recite, parrot) was measured (approximate onset to offset). A production in the mid range of each sample was chosen as the base duration. The LPC system at Bell Labs was used to lengthen or shorten the stressed syllable of the precursor word. Duration was changed by approximately 10% of the total duration of the base word. Thus, three versions of each sentence frame were created. The base durations and altered word durations are shown on Table 8.

The following measurements were done to determine the syllable durations for the target words. The base word was chosen from the mid range total duration of four target word productions. The duration of either the first syllable, the second syllable or a combination of both the first and second syllable of the base word was altered to create three versions of each target word: first syllable stress, second syllable stress and ambiguous stress.

The syllable durations for the ambiguous versions were calculated by determining the average proportional duration for each syllable (averaged across two stressed syllables and two unstressed syllables for each syllable). These proportions were used to calculate the appropriate duration of each syllable from the total duration of the base word. The syllable durations for the unambiguous version of the target words were calculated in a similar manner. Specifically, the average proportional duration of a stressed syllable from two productions of first syllable stress (for any one word) was used to calculate the appropriate duration for unambiguous first syllable stress. For example, the average proportional duration of the first syllable in the two productions of DIScount was 46%. The calculated duration for the first syllable to create an unambiguous first syllable stress was 209 msec (.46 x total duration of the base word).

Table 8

1. When you try to parrot **** quickly I will say it slowly.

    1000 msec   1350 msec

2. I think if you recite **** to me then I will know the word.

    1040 msec   1160 msec

Table 8. Sentence contexts used in experiment 2. Duration of context segments indicate in msec.

Table 9

Sentence frame	version	precursorword duration
parrot (bl)	base	355
(ll)	lengthened	390
(sl)	shortened	320
recite (b2)	base	446
(l2)	lengthened	496
(s2)	shortened	400

Table 9. Duration (msec) of percursor context word (parrot, recite) for each version (base, lengthened, shortened).

Similarly, the duration of the second syllable of this target word version was based on the average proportional duration of the second syllable of the same productions.

The duration of each syllable per version for each word are shown in Table 8. It should be noted that the calculated values were altered somewhat based on subjective judgements of the success of the duration change in obtaining the desired pattern. In addition, the procedure for the target words with vowel reduction (convict and conduct) differed somewhat. The first and second syllable version for each were based on different productions. Altering the duration of a single base proved to produce an unnatural utterance. Thus, a first syllable stress version was based on a first syllable production, a second syllable stress version was based on a second syllable stress production. The second syllable stress production was chosen to create the neutral version. Finally, sentence frames and target words were set at a steady 90 hz.

Each precursor context was combined with each word segment to create 18 conditions. These are shown on Table 9. The first four conditions were formed by combining an unambiguous target word with a lengthened or shortened context segment. These conditions had compatible sentence rhythm since the context timing information was

Table 10

	neutral stress (w0)	first stress (w1)	second stress (w2)
decrease	200 250	190 220	130 300
discount	210 250	240 250	190 300
conduct	200 240	230 220	175 270
convict	230 260	240 220	180 290

Table 10. Target durations (msec) used in experiment 2. The first number in each column is the duration of the first syllable. The second number in each column is the second syllable duration for each version.

appropriate for the target word. To create conditions with incompatible sentence rhythm, a context frame was combined with an unambiguous target word. In these conditions, the target word that occured was inappropriate to the timing of the context sentence.

Since each word also occured in an ambiguous form, an additional four condtions were created. These conditions consisted of a context that was compatible with either a stressed first syllable or a stressed second syllable. Sine the target word was ambiguous, these conditions were termed neutral-compatable and neutral-alternating.

In order to assess the influence of the duration manipulation in the context sentences, an additional six conditions were created . Each target word version was combined with the base sentence context. Two conditions consisted of combining a target word with a sentence context such that a _sw_ or a _ws_ pattern was not interrupted. Non-alternating conditions were created by combining a base context with a target word that interrupted the sentence stress pattern. The final two conditions combined the ambiguous word version with the base context to form neutral-alternation.

Table 11

combined segments	sentence rhythm	sentence stress
s2 w2 e2	COMP	ALT
l2 wl e2	COMP	NON-ALT
l1 wl el	COMP	ALT
sl w2 el	COMP	NON-ALT
l1 w2 el	INC	NON-ALT
ls wl el	INC	ALT
l2 w2 e2	INC	ALT
s2 wl e2	INC	NON-ALT
l1 w0 el	NEUT-COMP	NEUT-ALT
l2 w0 e2	NEUT-COMP	NEUT-ALT
sl w0 el	NEUT-COMP	NEUT-ALT
s2 w0 e2	NEUT-COMP	NEUT-ALT
bl wl el	NEUT	ALT
b2 w2 e2	NEUT	ALT
bl w2 el	NEUT	NON-ALT
b2 wl e2	NEUT	NON-ALT
bl w0 el	NEUT	NEUT-ALT
b2 w0 e2	NEUT	NEUT-ALT

Table 11. Conditions in experiment 2.

sl=shortened version of sentence 1 (sentences from table 6)
s2=shortened version of sentence 2
l1 & l2 = lengthened version
bl & b2 = base version
wl, w2& w0 = first, second andneutral word accent,respectively
el & e2 = end segment

Procedure The sentences were played on a Revox A77 tape recorder at a comfortable listening level over Layfayette SP-55 stereo headphones. As in experiment 1, subjects were asked to listen to each sentence and indicate perceived stress of the target item. The target item was identified for each sentence on an answer sheet. Subjects indicated lexical stress by underlining the appropriate syllable on the answer sheet. Subjects also indicated a rating response for each item in the space provided on the answer sheet.

Results

As in Experiment 1, a separate ANOVA was performed for each lexical item. The proportion of first syllable responses for all main effects and two way interactions are shown in Appendix C, tables 18 and 19, respectively. Appendix D, table 24 presents the values for individual conditions.

Decrease

Consider first the analysis for decrease. A significant main effect of target version ($F(2,48) = 14.26$, $p < .0001$) was found. The proportions of first syllable responses for the neutral and first syllable versions were quite similar (39% and 40%, respectively). A second syllable target was reported as a first syllable stress less frequently (20%). There was no effect of sentence stress ($F(1,42) < 1$) or sentence rhythm ($F(2,48)$ $p < 1$). Other significant effects included an order x sentence rhythm x target interaction ($F(4,168) = 2.31$, $p < .06$).

Discount

Main effects An ANOVA for the lexical item discount revealed three significant main effects. The main effect of target version ($F(1,42) = 10.36$, $p < .0001$) indicated the neutral and first syllable versions were perceived similarly (58% and 57%, respectively). This value was only 42% for the second syllable version. The sentence stress manipulation was significant for this lexical item ($F(1,42) = 3.38$, $p < .07$). A greater proportion of first syllable responses were reported for first syllable alternation than second syllable alternation (56% and 49%, respectively). Finally, there was a significant main effect of sentence rhythm ($F(2,84) = 3.33$, $p < .04$). The neutral rhythm showed 47% of the responses were reported as first syllable stress. Both first and second syllable sentence rhythm showed a greater proportion of first syllable responses than the neutral version (53% and 57%, respectively). A Newman Keuls analysis of individual means found that first syllable sentence rhythm had significantely more first syllable responses that the neutral sentence rhythm ($p<.01$). However, second syllable sentence rhythm did not differ from neutral sentence rhythm nor did first and second syllable sentence rhythm differ from each other.

Interactions Two interactions were significant : order x sentence rhythm ($F(2,84) = 2.86$, $p < .01$) and order x target ($F(2,84) = 4.23$, $p < .01$).

Convict

Main effects An ANOVA for the lexical item convict revealed two main effects. A significant target version effect ($F(2,84) = 334.21$, $p <.0001$) showed that a first syllable target was perceived as such (91%). The neutral and second syllable versions were rarely reported as first syllable stress (13% and 10%, respectively). Sentence stress was also significant ($F(21,42) = 4.43$, $p < .04$). Targets that appeared in a first syllable sentence stress context were reported as first syllable stress more frequently that when in a second syllable sentence stress (41% and 36%, respectively). Sentence rhythm was not significant ($F(2,84) < 1$).

Interactions A significant sentence stress x target interaction ($F(2,84) = 3.37$, $p < .02$) showed the following pattern of data (see Appendix C, table 19). Neutral and second syllable versions showed more first syllable responses if sentence stress favored a first syllable target relative to second syllable sentence stress (10% and 5%, respectively). The first syllable version showed slightly fewer first syllable responses in a first syllable sentence stress relative to a second syllable sentence stress (90% and 93%, respectively). Other significant effects included a sentence stress x sentence rhythm x target interaction ($F(4,168) = 2.71$, $p < .03$).

Conduct

Main effects Four main effects were found for the lexical item conduct (see Appendix C, table 18). The main effect of target version ($F(2,84) = 236$, $p < .0001$) revealed that neutral and second syllable versions tended to be reported as second syllable stress (11% and 7%, first syllable stress reports, respectively). The first syllable version tended to be reported as such (79%). Sentence stress reliably influenced reports ($F(1,42) = 3.25$, $p < .07$): first syllable sentence stress produced more first syllable responses than second syllable sentence stress (35% and 30%, respectively). In addition, a main effect of sentence rhythm ($F(2,84) = 5.38$, $p < .006$) revealed a first syllable sentence rhythm produced more fewer first syllable reports (35%) than second or neutral sentence rhythm (34% and 35%, respectively). A Newman-Keuls test found that the first syllable sentence rhythm produced significantly fewer first syllable responses than both neutral and second syllable rhythm ($p<.01$). Finally, a main effect of order was found ($F(1,42) = 7.18$, $p < .01$).

Interactions Five two way interactions were found (see Appendix C,table 19). Consider first the sentence stress x target interaction ($F(2,84) = 8.27$, $p < .001$). A first syllable target was reported as such to a greater extent if it appeared in a first syllable sentence stress relative to a second syllable sentence stress (87% and 72%, respectively). This effect was greatly reduced for the neutral version (12% and 9%, respectively) and slightly reversed for the second syllable version (5% and 9%, respectively). A significant sentence stress x sentence rhythm interaction ($F(2,84) = 6.36$, $p < .002$) showed fewer first syllable responses if both sentence stress and sentence rhythm favored a second syllable target (21%) than if neither or only one did (34% - 36%, see Appendix C, table 19).

Two additional two way interactions were significant: order x sentence rhythm ($F(2,84) = 10.46$, $p < .001$) and order x sentence stress ($F(1,42) = 6.77$, $p < .01$). Significant three way interactions included: order x sentence stress x target ($F(2,84) = 6.13$, $p < .003$), order x sentence rhythm x target ($F(4,168) = 6.33$, $p < .0001$) and sentence rhythm x sentence stress x target ($F(4,168)$ =

5.44, p < .0004). A single four way interaction was found: order x sentence rhythm x sentence stress x target (F(4,168) = 3.14, p < .01).

Rating Responses

Appendix D, table 25 presents the mean rating response for individual conditions. Identical ANOVAs were performed on the rating data for each lexical item. Summary data for main effects and two way interactions are shown on table 18 and 19, respectively.

Decrease

A main effect of target (F(2,84) = 22.43, p < .0001) was found for the lexical item decrease.

Discount

Comparable results were found for the target main effect for discount (F(2,84) = 11.37, p < .0001). An additional main effect of alternation was found (F(1,42) = 5.97, p < .01) and two way order x target interaction (F(2,84) = 5.67, p < .005). In general these lexical items (decrease and discount) were perceived as relatively ambiguous.

Convict

A main effect of target for the lexical item convict (F(2,84) = 2.54, p < .001) revealed this item was perceived as relatively unambiguous. Three interactions were found: order x sentence rhythm (F(2,84) = 6.85, p < .001), sentence stress x target (F(2,84) = 4.46, p < .01) and order x sentence rhythm x target (F(4,168) = 2.42, p < .05).

Conduct

A significant target effect for the lexical item conduct (F(2,84) = 194. 31, p < .0001) showed that this item was perceived as relatively unambiguous. Additional main effects included sentence stress (F(1,42) = 16.34, p < .0001), sentence rhythm (F(2,84) = 7.75, p < .001) and order (F(1,42) = 5.96, p < 01). Five two way interactions were found: sentence stress x sentence rhythm (F(2,84) = 9.98, p < .00001), sentence rhythm x target (f(4,168) = 5.41, p < .001) , order x sentence stress (F(1,42) = 11.08, p < .001) and order x sentece rhythm (F(2,84) = 11.32, p < .001). Additional interactions included: order x sentence stress x target (F(2,84) = 8.16, p < .0006), order x sentence rhythm x target (F(4,168) = 7.7, p < 10001), sentence stress x sentenc rhythm x target (F(4,168) = 7.01, p < .0001) and order x sentence stress x sentence rhythm x target (F(4,168) = 6.91, p < .0001).

Discussion

Experiment 1 was partially successful in replicating the effect of sentence stress found in experiment 1. Three of the four lexical items (DISCOUNT,CONVICT,CONDUCT) were influenced by sentence stress in the predicted direction. Sentence rhythm was statistically significant for two of the lexical items (DISCOUNT, CONDUCT). However, the interpretation of these data is somewhat problematic. The neutral sentence rhythm condition did not produce data that would allow meaningful comparison with the other sentence rhythm conditions. In one lexical item (CONDUCT) neutral and second syllable sentence rhythm produced the same proportion of responses. In addition, the first syllable sentence rhythm produced fewer first syllable responses than the other sentence rhythm conditions. In the case of the lexical item DISCOUNT, the neutral sentence rhythm condition produced fewer first syllable responses that either the first or second syllable sentence rhythm. Furthermore, the first first and second syllable rhythm conditions did not differ significantly from each other.

One possible explanation for the peculiar pattern of the sentence rhythm data may be found in the fact that normal variation in fundamental frequency was removed. As a result of this manipulation, the subjective quality of the speech was considerably reduced. The loss in intelligibility may have disrupted normal speech processes. Huggins (1978) has reported that intelligibility of speech is adversely affected by removal of fundamental frequency variation. Larkey (personal communication, 1983) has also found processing of sentences is disrupted if information provided by fundamental frequency is removed. Perhaps the relatively subtle changes involved in anticipatory timing were obscured by the unnaturalness of the stimuli.

A change in syllable duration as a cue for stress did not prove to be entirely successful. Although subjects tended to perceive stress as intended for first and second syllable stress, the proportion of responses and rating data indicated these tokens were perceived as relatively ambiguous. This suggests duration alone may not be a powerful indicator of stress. However, the quality of the speech in these stimuli makes this conclusion tentative. Resynthesis of the experimental sentences resulted in distortion of the speakers voice. This was simply due to an incompatability of the speaker's voice with computer resynthesis routines. Thus, there was intelligibility loss independent of interference from loss of pitch change information.

CHAPTER IV

GENERAL DISCUSSION

The experiments reported here investigated two alternative hypotheses concerning the role of rhythmic information on perception of stress in ambiguously stressed lexical items. The first hypothesis states that perception of stress is guided by an alternating pattern of stressed and unstressed syllables (sentence stress). Stress is reported such that this repeating pattern is not interrupted. The alternative hypothesis focuses on the relative timing of syllables (sentence rhythm). Huggins (1978) reported that the duration of stressed syllables is influenced by the surrounding rhythmic environment. The sentence rhythm hypothesis states that the listener uses these timing changes in lexical stress decisions. Clear evidence for sentence rhythm effects would suggest listeners have available quite detailed physical information. Small anticipatory changes in syllable duration could be used to constrain following rhythmic patterns.

Experiment one tested these hypotheses with naturally produced speech. Here two sentence stress patterns were produced with each lexical stress production. We assumed that violation of the alternating sentence stress pattern would result in rhythmic changes. Cross spliced sentences provided cases where sentence stress and sentence rhythm conflicted. Sentence stress was found to influence lexical stress reports in three of the four lexical items tested. However, no effect of sentence rhythm was found. Acoustic analysis of the context sentences revealed no rhythm induced changes in timing. Thus in order to convincingly demonstrate that sentence rhythm plays no role in lexical stress perception, a second experiment was performed.

In experiment two, linear predictive coding was used to alter naturally produced utterances. Here, the precursor word was lengthened (to anticipate a first syllable stress in the target item) or shortened (to anticipate second syllable stress). In addition, fundamental frequency was set to a constant value in an attempt to isolate duration changes from change in pitch. Again, the major influence on lexical stress reports was sentence stress. The role of sentence rhythm was not supported.

The two experiments thus found that sentence stress influenced lexical stress reports. The finding of sentence stress effects is consistent with a model of speech processing that incorporates a role of surrounding prosodic environment. These experiments suggest that once the listener extracts a prosodic pattern from the acoustic signal, this pattern is used even if contrary evidence is available in the physical signal. Although sentence rhythm information was provided in experiment 2, it was not utilized. Instead, listeners used the already established pattern. This suggest that sentence stress effects in experiment 1 (in which only sentence stress information was available) were not simply due to a default use of sentence stress information. The availability of sentence rhythm in experiment 2 did not alter use of the sentence stress pattern.

These findings support the notion proposed by Martin (1970) that once a rhythmic pattern is established, 'hearing' some stresses is based on a listeners knowledge of the language. One such type of knowledge may be that there is a tendency for stress shifts to occur to avoid adjacent stressed syllables in production. For example, the stressed syllable of the lexical item thirteen is shifted to the first syllable when produced in the context thirteen men. Of course, two adjacent stressed syllables may occur in naturally spoken language. In the present experiments, relatively unambiguous lexical stress reduced the influence of sentence stress. In these cases, listeners did report lexical stress such that two stressed syllables occur adjacently. This suggests that lexical stress assignment involves an interaction of context stress patterning and segmental information (such as vowel quality). However, in the absence of clear acoustic information, lexical stress pattern is determined by the prosodic context (sentence stress).

The lack of acoustic anticipatory timing changes in the sentence frames used in experiment 1, suggest a possible relation between nuclear accent and rhythmic anticipation. Nuclear accent typically falls on the final word in a syntactic clause. Active anticipation via rhythmic timing would be misleading at these points since rhythmic changes do not cross major syntactic boundaries. Phrase final lengthening occurs regardless of the stress pattern of the following phrase. Nuclear accent may fall in phrase final position as a default in order to maintain informative rhythmic changes.

In order to explain why the sentence stress pattern is maintained by listeners, it is necessary to consider the role of stress in language processing. It is well known that stressed segments are physically more informative than unstressed segments: the acoustic signal is clearer. Spoken language processing would be facilitated if attentional capacity was devoted to informative segments (cf, Martin, 1975; Cutler, 1979). Knowledge of rhythmic constraints would allow adequate anticipation of points of clarity.

Stressed syllables may also provide word boundary information. Taft (1983) found that listeners tend to segment lexically ambiguous strings such that a new word begins with a stressed syllable. If prosody constrains the occurrence of stressed syllables then this may be early information for a possible word boundary.

The conclusions of the present experiments may only cautiously be extended to natural speech. The pattern of stressed syllables in the experimental sentences was somewhat artificial. It is not generally the case that one finds in natural speech such a regular pattern of alternating syllables. Further experiments would be necessary to establish the generality of the influence of sentence stress on reports of lexical stress. Sentence contexts could be constructed with more natural stress patterns.

A second question concerns the locus of the effect. Two alternative models suggest themselves. In one case, the locus of the effect is perceptual. The patterning of stressed and unstressed syllables is extracted from th physical signal. This knowledge is used very early to make decisions concerning lexical stresss. Perception of lexical stress is based on this established pattern. This model is consistent with current psychological (cf. Martin, 1972) and linguistic (cf. Thompson, 1980) proposals concerning sentence rhythm and lexical stress. A second possibility is that lexical stress is reconstructed based on a memory representation of the stress patterning of the context. In the present experiments, subjects responded after each sentence was completed. One way to explore the level at which sentence prosody has its effect would be to interrupt presentation of the sentence immediately after the lexical item or further downstream. The perceptual hypothesis would be supported if it was the case that an immediate speeded response produced the same magnitude of effect as an end of sentence response.

Finally, listeners in these experiments had to rely solely on prosodic information (and/or vowel quality of the target) to make lexical stress decisions. The sentences were constructed such that sentence structure was neutral in terms of the grammatical role of the lexical item. In natural speech, it is typically the case that verbs receive second syllable stress while nouns receive first syllable stress in their citation form. It would be of interest to determine if syntactic category influenced lexical stress decisions in a fashion similar to sentence stress. Consider the set of sentences listed in table 12. Sentences 1 and 3 are sentences in which the noun form of the word <u>permit</u> is used. Sentences 2 and 4 are cases where the verb form is used. Notice that the sentence stress pattern in sentences 1 and 4 is compatable with the canonical stress pattern of the lexical item permit. Sentences 2 and 3 are cases in which the grammatical category , and thus the canonical stress pattern, of permit is not compatable with sentence stress. If listeners are influenced by the syntactic category of an ambiguously stressed lexical item, then judgements of lexical stress in sentence 2 would produce more verb responses than sentence 1. Similarly, judgements of lexical stress in sentence 4 would produce more verb responses than in sentence 3. An opposite pattern of results would hold if sentence stress was more influential. It may be the case that different subject strategies would hold depending on the task (cf. Miller, Green and Schermer, 1982). Subjects required to attend to the meaning of the sentences may be primarily influenced by syntactic variables. However, if a perceptual orientation was primary, then sentence stress may be more influential. Sentence stress effects in both task orientations would support a perceptual model. A model in which sentence stress operates at a post perceptual level would be supported if sentence stress effects were obtained only in a perceptual task.

Table 12

1. In the southern states,
s w s w s

a drivers permit must be granted.
w s w s w s w

2. On the trip,
s w s

the drivers permit kids to sing some songs.
w s w s w s w s

3. In northern Maine,
w s s

a court permit requires much detail.
w s w s w s w

4. In northern Maine,
w s w s

the courts permit thevoters little privacy.
w s w s w s w s w

Reference Notes

1. The role of sentence stress and attentional capacity is not entirely clear cut. Cutler and Fodor (1979) havefound faster phoneme reaction time to targets that are semantically focused but do not receive sentence stress. This indicatesthatphoneme reaction time facilitationin sentence stress position is at least partly a function of semantic focus.

REFERENCES

Bolinger, D. (1964) Around the edge of language: intonation. Harvard Educational Review, 34, 282-296.

Cooper, W. E., Paccia, J. M., & Lapointe, S. G. (1978) Hierarchical coding in speech timing. Cognitive Psychology, 10, 154-177.

Cutler, A. (1976) Phoneme monitoring RT as a function of preceding intonation contour. Perception & Psychophysics, 20, 55-60.

Cutler, A. (1979) Intonation, timing and sentence comprehension: A review of the psycholinguistic evidence. Paper presented at the conference of the Mental Representation of Phonology, Amherst, MA.

Cutler,A., & Clifton, C. The use of prosodic information in word recognition. Attention and Performance, X, Eindhoven, July 1982.

Cutler,A., & Fodor, J. (1979) Semantic focus and sentence comprehension. Cognition ,7,49-59.

Cutler, A., & Foss, D. J. (1977) On the role of sentence stress on sentence processing. Language and Speech, 20, 1-10.

Darwin, C. J. (1975) On the dynamic use of prosody in speech perception. In A. Cohen and S.G. Nooteboom (Eds.) Structure and Process in Speech Perception: Proceedings of the Symposium on Dynamic Aspects of Speech Perception, I.P.O., Eindhoven, Netherlands., 161-176.

Delattre, P. C., Liberman, A. M., & Cooper, F. S. (1955) Acoustic loci and transitional cues for consonants. Journal of the Acoustical Society of America , 27, 769-773.

Donavan, A., & Darwin, D. J. (1979) The perceived rhythm of speech. Ninth International Congress of Phonetic Sciences, Copenhagen.

Dooling, D. J. (1974) Rhythm and syntax in sentence perception. Journal of Verbal Learning and Behavior , 13, 255-383.

Forster, K. (1976) Accessing the mental lexicon. In R. J. Wales and E. Walker (Eds.), New Approaches to Language Mechanisms , Amsterdam: North-Holland Press, 1976, 257-288.

Forster, K. I. (1979) Levels of processing and the structure of the language processor. In W.E. Cooper and E.T. Walker (Eds.),Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett , Lawrence Erlbaum, 27-85.

Fry, D. B. (1958) Experiments in the perception of speech. Language and Speech , 1, 120-152.

Gay, T. (1978) Physiological and acoustic correlates of perceived stress. Language and Speech , 21, 347-353.

Huggins, A. W. (1978) Speech timing and intelligibility. In J. Requin (Ed.), Attention and Performance, VII , Hillsdale,N.J.: Erlbaum.

Huss, V. (1978) English word stress in post nuclear position. Phonetica , 35, 86-105.

Klatt, D., & Cooper, W. E. (1975) Perception of segment duration in sentence contexts. In A. Cohen and S. G. Nooteboom (Eds.), Structure and process in speech perception , Heidelberg, Germany: Springer-Verlag.

Lehiste, I. (1973) Rhythmic units and syntactic units in production and perception. Journal of the Acoustical Society of America , 54, 1228-34.

Lehiste, I., Olive, J., & Streeter, L. (1976) Role of duration in disambiguating syntactically ambiguous sentences. Journal of the Acoustical Society of America, 60, 1199-1202.

Levinson, S.C. (1983) Pragmatics, Cambridge University Press.

Lyons, J. (1977) Semantics , Cambridge University Press.

Marslen-Wilson, W. D., & Welsh, A. (1978) Processing interactions and lexical access during word recognition in continuous speech. Cognitive Psychology , 10, 26-63.

Marslen-Wilson, W. D. (1975) Sentence perception as an interactive parallel process. Science , 189, 226-228.

Martin, J. G. (1970) Rhythm induced judgements of word stress in sentences. Journal of Verbal Learning and Verbal Behavior , 9, 451-454.

Martin, J. G. (1972) Rhythmic vs. serial structure in speech and other behavior. Psychological Review, 79, 487-509.

Martin, J. G. (1975) Rhythmic expectancy in continuous speech perception. In A.Cohen and S. G. Nooteboom (Eds.) Structure and Process in Speech Perception: Proceedings of the Symposium on Dynamic Aspects of Speech Perception, 161-176.

Martin, J. G. (1979) Rhythmic and segmental perception are not independent. Journal of the Acoustical Society of America , 65(5), 1268-1297.

Miller, J. L., Green, K., & Schermer,T. On the distinction between prosodic and semantic factors in word identification. Paper presented at Acoustical Society of America Meeting, April 26-30, 1982, Chicago.

Miller, J. L., & Grosjean, F. How the components of speaking rate influence perception of the phonetic segments. Journal of Experimental Psychology: Human Perception and Performance, in press.

Nakatani, L., & Aston, C. H. (1978) Perceiving stress patterns of words in sentences. Journal of the Acoustical Society of America , 63, S55 (abstract).

Nakatani, L., & Shaffer, J. (1978) Hearing "words" without words: prosodic cues for word perception. Journal of the Acoustical Society , 63, 234-245.

Oshika, B. T., Zue, V. W., Weeks, R. V., Nue, H., & Aurbach, J. (1975) The role of phonological rules in speech understanding research. IEEE Transactions, Speech Signal Processing , ASSP-23, 104-112.

Pierrehumbert, J. (1980) The phonology and phonetics of English intonation . Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Port, R. F. (1978) The influence of word-internal versus word-external tempo on the voicing boundary for medial stop closure. Journal of the Acoustical Society of America , 63, S20 (abstract).

Shields, J. L., McHugh, A., & Martin, J. G. (1974) Reaction time to phoneme targets as a function of rhythmic cues in continuous speech. _Journal of Experimental Psychology_ , 102, 250-255.

Slowiaczek, M. L. (1981) _Prosodic units as language processing units_ . Unpublished doctoral disseration, University of Massachusetts at Amherst, Amherst, MA.

Taft, L. _A Study of the Perception of Stress in Initial Syllables_. Unpublished manuscipt, May 1980.

Taft, L. (1983) _Prosodic constraints and lexical segmentation_ . Unpublished doctoral dissertation, University of Massachusetts at Amherst, Amherst, MA.

Thompson, H. S. (1980) _Stress and salience in English: theory and practise_. Palo Alto Research Center, CSL-80-8.

APPENDIX A

Table 13

	Lexical item			
	DECREASE	DISCOUNT	CONVICT	CONDUCT
Sentence stress				
first	43	57	52	(51)
second	33	43	46	(50)
Sentence Rhythm				
first	(38)	(52)	(50)	(51)
second	(39)	(47)	(49)	(51)
Target				
first	48	59	88	91
second	28	40	11	9

Table 13. Percentage of first syllable responses for main effects of sentence stress, sentence rhythm and target. Note--non significant effects indicated by parenthesis. Experiment 1.

Table 14

	Lexical item							
	DECREASE		DISCOUNT		CONVICT		CONDUCT	
	Target							
	1	2	1	2	1	2	1	2
Sentence Stress								
first	57	28	(67)	(48)	88	17	(45)	(55)
second	38	27	(51)	(32)	89	05	(48)	(51)
Exaggeration								
YES	31	27	55	44	(89)	(87)	(87)	(95)
NO	64	29	63	35	(13)	(09)	(06)	(11)

Table 14. Percentage of first syllable responses for two way interactions. Note--non-significant effects indicated by parenthesis. Experiment 1.

Table 15

	Lexical Item			
	DECREASE	DISCOUNT	CONVICT	CONDUCT
Sentence Stress				
first	3.74	3.25	(3.49)	(3.47)
second	4.02	3.78	(3.48)	(3.52)
Target				
first	3.4	3.13	1.89	1.86
second	4.27	3.91	5.07	5.14
Exaggeration				
NO	4.2	(3.52)	(3.49)	3.66
YES	3.5	(3.55)	(3.48)	3.34

Table 15. Rating response for main effects. Experiment 1.

Table 16

	Lexical Item			
	DECREASE		DISCOUNT	
	Target			
	1	2	1	2
Exaggeration				
YES	4.09	4.21	3.28	3.76
NO	2.89	4.32	2.88	4.65
Sentence Rhythm				
First	3.44	4.37	2.90	4.00
Second	3.54	4.17	3.20	3.82

Table 16. Mean rating response for significant two way interactions. Experiment 1.

APPENDIX B

Table 17

DECREASE	Phrase Duration	Context Duration	Proportion due to context
wsws	889	479	.54
wssw	894	474	.53
swsw	743	363	.49
swws	682	342	.50
DISCOUNT			
wsws	866	443	.49
wssw	892	448	.50
swsw	786	365	.46
swws	798	328	.41
CONVICT			
wsws	872	472	.54
wssw	945	475	.50
swsw	773	353	.46
swws	829	372	.45
CONDUCT			
wsws	876	466	.53
wssw	950	470	.49
swsw	879	389	.44
swws	839	369	.44

Table 17. Context word duration (msec) proportional tc total duration of target word and context word, experiment 1.

APPENDIX C

Table 18

	Lexical item			
	DECREASE	DISCOUNT	CONVICT	CONDUCT
Sentence Stress				
first	(33)	56	41	35
second	(33)	49	36	30
Sentence Rhythm				
first	(34)	57	(37)	28
second	(33)	53	(40)	35
neutral	(32)	47	(37)	34
Target				
first	39	58	91	79
second	20	42	10	11
neutral	40	57	13	7

Table 18. Percentage of first syllable responses for main effects of sentence stress, sentence rhythm and target. Note--non-significant effects are indicated by parenthesis. Experiment 2.

Table 19

	CONDUCT			CONVICT		
	1	2	neut	1	2	neut
Sentence Stress						
first	87	5	12	90	12	18
second	72	9	9	93	7	8

Lexical item; Target

Sentence Rhythm	1	2	neut
first	85	10	10
second	65	5	13
neutral	88	6	9

Sentence Rhythm	Sentence stress 1	Sentence stress 2
first	36	34
second	35	21
neutral	33	36

Table 19. Percentage of first syllable response for significant two way interactions. Note--non-significant results indicated by parentheses. Experiment 2.

Table 20

	Lexical item			
	DECREASE	DISCOUNT	CONVICT	CONDUCT
Sentence Stress				
first	(4.03)	3.22	3.83	4.04
second	(4.10)	3.52	4.07	4.38
Sentence Rhythm				
first	(4.02)	(3.24)	(4.03)	(4.44)
second	(4.03)	(3.40)	(3.90)	(4.08)
neutral	(4.15)	(3.46)	(3.91)	(4.10)
Target				
first	3.75	3.16	1.60	2.21
second	4.62	3.74	5.19	5.30
neutral	3.83	3.20	5.05	5.12

Table 20. Rating responses for main effects. Note--non-significant effects indicated by parenthesis. Experiment 2.

Table 21

	Lexical item					
	CONVICT			CONDUCT		
	Target					
	1	2	neut	1	2	neut
Sentence Stress						
first	1.80	5.30	5.01	1.65	5.04	4.79
second	2.62	5.30	5.22	1.55	5.34	5.31
Sentence Rhythm						
first	2.82	5.36	5.14			
second	1.90	5.20	5.14			
neutral	1.92	5.34	5.06			

	Sentence stress	
	1	2
Sentence Rhythm		
first	4.07	4.81
second	3.98	4.18
neutral	4.06	4.15

Table 21. Rating response syllable response for significant two way interactions. Experiment 2.

APPENDIX D

Table 22

	TARGET							
	first				second			
	EXAGGERATION							
	yes		no		yes		no	
	Sentence Rhythm							
	1	2	1	2	1	2	1	2
DECREASE								
Sentence Stress								
first	.38	.36	.79	.77	.29	.34	.29	.22
second	.22	.29	.50	.52	.20	.25	.29	.36
DISCOUNT								
Sentence Stress								
first	.61	.63	.75	.68	.54	.43	.45	.50
second	.63	.31	.56	.54	.38	.43	.25	.22
CONVICT								
Sentence Stress								
first	.84	.90	.86	.90	.27	.13	.09	.18
second	.88	.93	.93	.79	.06	.04	.04	.04
CONDUCT								
Sentence Stress								
first	.86	.86	1.0	1.0	.04	.04	.13	.09
second	.90	.86	.90	.90	.09	.09	.04	.18

Table 22. Summary table of percentage of first syllable responses, experiment 1.

Table 23

	TARGET							
	first				second			
EXAGGERATION	yes		no		yes		no	
SENTENCE RHYTHM	1	2	1	2	1	2	1	2
DECREASE								
Sentence Stress								
first	3.90	4.00	2.29	2.68	4.25	4.06	4.38	4.34
second	4.22	4.25	3.36	3.25	4.56	4.43	4.29	3.84
DISCOUNT								
Sentence Stress								
first	2.90	3.11	2.68	2.68	3.52	3.70	3.84	3.59
second	2.80	4.27	3.36	3.22	4.00	3.84	4.63	4.15
CONVICT								
Sentence Stress								
first	2.11	1.68	1.90	1.81	4.72	5.11	5.20	4.75
second	2.04	1.70	1.77	2.11	5.36	5.18	5.29	4.97
CONDUCT								
Sentence Stress								
first	1.81	2.22	1.54	1.54	5.54	5.27	4.86	5.00
second	2.09	1.77	2.04	1.86	5.22	5.36	5.04	4.81

Table 23. Summary table of rating responses, experiment 1.

Table 24

	TARGET first			TARGET second			Neutral		
SENTENCE RHYTHM	1	2	N	1	2	N	1	2	N
DECREASE									
Sentence Stress first	.29	.47	.47	.25	.18	.20	.36	.38	.34
second	.43	.38	.34	.27	.15	.15	.45	.40	.40
DISCOUNT									
Sentence Stress first	.65	.61	.61	.52	.45	.36	.70	.56	.56
second	.54	.56	.45	.45	.38	.36	.56	.59	.50
CONVICT									
Sentence Stress first	.95	.90	.84	.09	.15	.13	.13	.22	.20
second	.86	.93	1.00	.09	.09	.04	.11	.09	.04
CONDUCT									
Sentence Stress first	.86	.86	.88	.02	.11	.02	.18	.11	.09
second	.45	.84	.88	.09	.09	.11	.09	.09	.09

Table 24. Summary table of percentage of first syllable responses, experiment 2.

Table 25

	TARGET first			TARGET second			Neutral		
	SENTENCE RHYTHM 1	2	N	1	2	N	1	2	N
DECREASE									
Sentence Stress first	3.93	3.47	3.47	4.40	4.54	4.63	4.00	3.79	3.97
second	3.75	3.70	4.06	4.43	4.86	4.88	3.36	3.84	3.77
DISCOUNT									
Sentence Stress first	2.81	2.97	2.93	3.52	3.72	3.59	2.86	3.38	3.18
second	3.40	3.18	3.65	3.54	4.02	4.06	3.29	3.13	3.38
CONVICT									
Sentence Stress first	1.52	1.61	1.81	5.31	5.00	4.81	4.93	4.72	4.72
second	1.79	1.52	1.34	5.45	5.27	5.31	5.20	5.27	5.47
CONDUCT									
Sentence Stress first	1.86	1.75	1.81	5.36	5.06	5.47	5.00	5.13	4.90
second	3.79	2.06	2.02	5.36	5.34	5.20	5.29	5.15	5.22

Table 25. Summary table of rating responses, experiment 2.