Working Papers of the Cornell Phonetics Laboratory
1993, vol. 8, pp. 119-140.

Acoustic correlates of form class

Joan A. Sereno and Allard Jongman

1. Introduction

A number of recent experimental studies have begun to demonstrate the importance of lexical stress cues for segmentation and syllabification strategies (e.g., Treiman, 1989; Cutler, 1991). Stress information also plays a key role in many recent theories of phonological structure both at lexical (Chomsky and Halle, 1968; Hayes, 1981; Liberman and Prince, 1977) and sentential levels (Beckman, 1986; Pierrehumbert, 1980). These studies convincingly argue that lexical stress information is crucial for the processing of continuous speech.

Four relevant acoustic parameters have been identified as possible cues to lexical stress. The present research uses these acoustic cues to investigate grammatical class differences in English. The question addressed is whether there are any systematic acoustic differences that can distinguish grammatical classes (i.e., nouns and verbs) in English.

Beginning in the 1950's, a number of studies investigated the acoustic correlates of lexical stress in a variety of languages including English, Polish, French, and Swedish (for reviews, see Gay, 1978; Lehiste, 1970). These studies concentrated on four acoustic measures of perceived stress: duration, intensity, fundamental frequency, and spectral composition. In general, longer duration, greater amplitude, higher fundamental frequency, and less vowel reduction in a syllable contribute to the perception of stress (Bolinger, 1958; Fry, 1955, 1958; Lieberman, 1960; Lindblom, 1963). However, the individual contribution of each of these factors in signaling lexical stress remains unclear. While some studies find that fundamental frequency appears to be the most predominant cue to perceived stress, variations in duration, amplitude and formant structure also systematically contribute to stress judgments. Moreover, the relative importance of each of these parameters varies with the position of the lexical item in the sentence, suggesting a number of interactions (Morton and Jassem, 1965; Gay, 1978; Nakatani and Aston, 1978) In speech production, then, a complex of acoustic cues, including fundamental frequency, duration, intensity, and spectral composition, appears to collectively contribute to the perception of contrastive stress.

2. Lexical statistics

Many studies examining the acoustic properties characterizing stress have concentrated on a rather limited set of bisyllabic, stress-contrastive words in English. For these bisyllabic words, a shift in stress from the first to the second syllable is associated with a shift of grammatical class. For example, 'survey' is stressed on the first syllable when used as a noun (*They conducted a SURvey of the population*), but stressed on the second syllable when used as a verb (*They must surVEY the land*). The use of such words allows for a constant segmental content.

The characteristic stress pattern in which nouns are forestressed and verbs are backstressed is quite pervasive in English. Although the linguistic literature commonly references this observation (e.g., Chomsky and Halle, 1968; Sherman, 1975; Liberman and Prince, 1977), few attempts were made to statistically quantify the pattern.

Recently, however, a number of studies have begun to systematically investigate this striking noun-verb stress asymmetry (Cutler and Clifton, 1984; Kelly, 1988; Kelly and Bock, 1988; Sereno, 1986). In a lexical analysis of English, Sereno (1986) demonstrated that most bisyllabic nouns in English are forestressed whereas most bisyllabic English verbs are backstressed. Table 1 shows a breakdown of the Brown Corpus (Francis and Kucera, 1982) in terms of number of syllables for pure nouns (words used only as nouns in English), pure verbs (words used only as verbs in English), and ambiguous items (words with both noun and verb instances in English). Lexical items included words with a frequency of 5 per million or greater. For each item, the number of syllables was established by its transcription in *Webster's Seventh New Collegiate Dictionary* (1963).

NOUNS (Total number = 3,858)

number of syllables

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
536	1,425	1,061	586	210	35	3	2
(14)	(37)	(28)	(15)	(5)	(1)	(0)	(0)

VERBS (Total number = 999)

number of syllables

1	**2**	**3**	**4**	**5**
187	523	222	62	5
(19)	(52)	(22)	(6)	(1)

NOUN/VERB 'AMBIGUOUS' (Total number = 1,694)

number of syllables			
1	**2**	**3**	**4**
1,014	570	97	13
(60)	(34)	(6)	(1)

Table 1. A lexical analysis of the Brown Corpus (Francis and Kucera, 1982) showing the number of syllables for nouns, verbs, and noun/verb ambiguous words in English. Numbers in parentheses represent percentages.

Sereno (1986) found that 93% of the 1,425 bisyllabic nouns were forestressed while 76% of the 523 bisyllabic verbs were backstressed. A similar relation between stress and grammatical class was reported by Kelly and Bock (1988). These data, of course, provide the possible motivation for that select set of English words that change stress pattern depending on their grammatical class.

An analysis of the noun/verb ambiguous items was also conducted (Sereno, 1986). Table 2 provides a breakdown of the 570 bisyllabic ambiguous words. Stimuli are analyzed in terms of location of the stressed syllable and frequency dominance of the grammatical category. The words are categorized as either being forestressed, backstressed, or variably stressed. Variable stress indicates words that allow alternate stress placement. Dominance is characterized as noun dominant, verb dominant, or equi-dominant. A noun or verb dominant word is a word which occurs more than 50% of the time as a noun or verb, respectively, while an equi-dominant word is used exactly 50% of the time as a noun and 50% as a verb.

BISYLLABIC NOUN/VERB 'AMBIGUOUS' WORDS

Stress Location	**Noun Dominant**	**Verb Dominant**	**Equi-Dominant**	**Total**
Forestress	293 (75)	85 (22)	14 (3)	392
Backstress	66 (55)	51 (43)	2 (2)	119
Variable	31 (52)	27 (46)	1 (2)	59
				570

Table 2. Analysis of bisyllabic noun/verb ambiguous words in terms of stress placement (forestress, backstress, and variable stress) and grammatical category dominance (noun dominant, verb dominant, and equi-dominant).

The variable items consist of the small minority (10%) of bisyllabic ambiguous words in which a change of grammatical class from noun to verb produces a shift in stress from the first to the second syllable (for a list of words, see Quirk, Greenbaum, Leech, and Svartvik, 1986). However, the vast majority (90%) of this ambiguous group includes words that maintain a constant stress pattern across syntactic function (e.g., *answer*, *design*.) These ambiguous stimuli do not shift stress placement with grammatical class membership. Nevertheless, these words show consistent stress placement depending on their dominant grammatical class usage. That is, for these stimuli, there is a significant interaction between stress placement and dominance [$\chi^2 = 19.97$, $p < .001$]. For these ambiguous stimuli, forestressed stimuli are used relatively more often as nouns (75%) than verbs (22%) while backstressed stimuli are used more equally often as verbs (43%) and as nouns (55%).

The tendency for nouns to be forestressed and verbs backstressed also shows up in production studies. Recent findings by Kelly and colleagues (see Kelly, 1992, for a review) show that subjects, when asked to read aloud sentences containing bisyllabic nonwords, produced these nonwords more often with first-syllable stress when the syntactic context indicated that the nonword functioned as a noun, and with second-syllable stress when the context indicated a verb (Kelly and Bock, 1988). In a subsequent study, Kelly (1988) showed that subjects in a sentence production task used forestressed nonwords more often as nouns and backstressed nonwords more often as verbs.

Sensitivity to the noun-verb stress difference is also suggested by a recent on-line study by Kawamoto, Farrah, and Overbeek (1992). Although Kawamoto et al. (1992) found that naming latencies to forestressed or backstressed words were not differentially affected by grammatical category (noun or verb), they did show a consistent effect when stimuli were presented in context. That is, when preceded by an appropriate syntactic context, words with stress patterns that were consistent with their grammatical category (i.e., first-syllable stress for nouns and second-syllable stress for verbs) were named more quickly while words with stress patterns inconsistent with their grammatical category were named more slowly. Taken together, these findings provide evidence for the claim that speakers' knowledge of the systematic relation between grammatical category and stress is operative in speech production.

Interestingly, Cutler and Clifton (1984) have shown that the systematic relation between grammatical category and stress pattern in English may not be used in perception. In a word recognition experiment, a one-word context predicting the part of speech of the following noun or verb did not speed recognition to canonically stressed words (i.e.,

forestressed nouns, backstressed verbs). This result suggests that listeners did not use stress information as an aid in the identification of individual words. However, in a second experiment, Cutler and Clifton (1984) did show that misstressed words were more difficult to recognize, suggesting that at least erroneous stress information affects the recognition process.

Acoustically, the noun-verb stress difference is traditionally illustrated with noun-verb homographs such as the pair *SURvey* (noun) and *surVEY* (verb). Such lexical items shift stress placement with a change in grammatical class and show substantial acoustic differences (e.g., Gay, 1978). The vast majority (90%) of categorically ambiguous bisyllabic words, however, do not shift stress with a change in grammatical class (e.g., *answer, design*). The goal of the present study was to discover whether systematic acoustic differences could also be observed in these words which do not exhibit contrastive stress changes. In the present experiment, ambiguous lexical items which change grammatical class but not stress placement were analyzed for acoustic differences. If an underlying pattern of stress assignment based on grammatical class usage is operative in speech production, then slight, stress-related acoustic differences in grammatically ambiguous words read in two different grammatical class contexts may be detectable.

3. Methods

3.1 Subjects. Five Brown University undergraduates (two male, three female) participated as paid volunteers in this experiment. Speakers 1 and 2 are male and Speakers 3, 4 and 5 are female. All were native speakers of American English with no known history of speech or hearing impairment. The speakers had no training in linguistics.

3.2 Stimuli. The recorded utterances were 16 bisyllabic words produced by the five speakers (See Table 3). The words were recorded on magnetic tape in a sound-treated room with a Nagra 4.2 tape recorder and a Shure SM 81 microphone at the Brown University Phonetics Laboratory.

The word stimuli were all ambiguous with respect to grammatical class in the sense that they could be used either as nouns or verbs without change in lexical stress. Stimulus items were balanced both with respect to form class frequency as well as stress assignment. For one half of the stimuli, the frequency of noun usage was greater than verb usage, while for the other half, the frequency of verb usage was greater than noun usage. In addition, one half of noun-dominant words and one half of verb-dominant words were stressed on

the first syllable while the remaining words were stressed on the second syllable. Consequently, there were four stimulus items in each of the following groups: noun-dominant and forestressed (e.g., *poison*; mean noun usage frequency 66; mean verb frequency usage 33); noun-dominant and backstressed (e.g., *debate*; mean noun usage frequency 125; mean verb frequency usage 76); verb-dominant and forestressed (e.g., *notice*; mean noun usage frequency 21; mean verb frequency usage 53); and verb-dominant and backstressed (e.g., *escape*; mean noun usage frequency 18; mean verb frequency usage 48).

Noun-dominant First-Syllable Stress	Noun Frequency	Verb Frequency
favor	63	49
poison	11	7
practice	132	40
struggle	57	36
Noun-dominant Second-Syllable Stress	Noun Frequency	Verb Frequency
control	220	95
debate	36	10
dispute	37	14
report	205	184
Verb-dominant First-Syllable Stress	Noun Frequency	Verb Frequency
handle	22	81
notice	39	84
rescue	9	14
welcome	13	33
Verb-dominant Second-Syllable Stress	Noun Frequency	Verb Frequency
embrace	6	18
escape	24	69
neglect	8	28
reply	35	75

Table 3. The 16 stimulus words listed with their frequency of occurrence as nouns and verbs.

3.3 Procedure. The 16 stimuli were read and recorded in both a noun context (noun reading) and in a verb context (verb reading). The noun context consisted of a list of 75 multisyllabic pure nouns (words used only as nouns in English) and the verb context consisted of a list of 75 multisyllabic pure verbs (words used only as verbs in English). The 16 stimuli were then interspersed in the noun and verb context lists. Each of the 16 ambiguous stimulus words was separated by at least four unambiguous lexical items in each list. The order of the contexts in which the stimulus words were read was counterbalanced across speakers. Three speakers read the stimuli in the noun context first and the verb context second while two speakers read the stimuli in the verb context first and the noun context second. Between readings, speakers participated in an unrelated perception experiment. When questioned after completion of the recording, none of the speakers had noticed that some words occurred twice in the lists. Two instances of each stimulus word were then analyzed, yielding a total of 32 stimulus items (16 noun reading stimuli and 16 verb reading stimuli) for each speaker.

3.4 Acoustic Analysis. Stimuli were analyzed at the Max Planck Institute for Psycholinguistics. Stimuli were digitized at 10 KHz on a VAX 740 computer and low-passed filtered at 4.5 kHz . In order to avoid bias while analyzing the stimuli, all test words were coded numerically so that the experimenters did not know whether they were dealing with the noun or verb reading of a particular stimulus item. All speech stimuli were displayed on a graphics terminal. First, the onset and offset of the stimuli were identified, and the boundary between the first and second syllables established on the basis of both auditory and visual examination of the stimuli. The criteria used in determining the onset of the second syllable were straightforward. For half of the stimuli, the onset of the second syllable was identified as the onset or closure of the first medial consonant which was not part of the onset cluster (*favor, poison, struggle, debate, report, notice, neglect, reply*). For the other half of the stimuli, the onset of the second syllable was identified as the closure for the second consonant of the medial cluster (*practice, control, dispute, handle, rescue, welcome, embrace, escape*). To ensure systematic comparisons between stimuli, identical criteria for both the noun and verb reading of a particular stimulus item were used.

In order to control for variations in speaking rate, intensity, or fundamental frequency between recordings of the two lists, a ratio of the first to the second syllable was used for all measurements (see, for example, Fry 1955; 1958). A valid comparison of ratios in each of two readings of a stimulus word could thus be secured regardless of the idiosyncratic circumstances under which the words were produced.

Three acoustic measurements of the stimuli (duration, amplitude, and fundamental frequency) were obtained. Absolute duration values were calculated for both syllables of each stimulus item, and a ratio of first-to-second syllable duration was computed for each item. RMS-amplitude was measured by pitch-synchronously placing a full Hamming window over each syllable of the stimulus words. Since a change in a subject's overall loudness affects the absolute dB level, a ratio of first-to-second syllable amplitude was computed for each item. Finally, fundamental frequency (F0) contours were obtained for each syllable, using a pitch extraction algorithm based on an autocorrelation procedure (Reetz, 1989). Since F0 contours for the stimulus items were quite flat and a visual examination revealed little disparity between noun and verb readings of a particular word, a single value (average F0) per syllable was used. A ratio of first-to-second syllable fundamental frequency was then computed for each item.

4. Results

For each parameter under investigation (i.e., duration, amplitude, fundamental frequency), a three-way ANOVA was conducted with the factors Stress (Forestressed or Backstressed), Dominance (Noun Dominant or Verb Dominant), and Reading (Noun Reading or Verb Reading). The complete data for each parameter under investigation are presented in Appendix A. Table A1 gives the mean duration ratio values, Table A2 the mean amplitude ratio values, and Table A3 the mean fundamental frequency ratio values. All data are shown for each subject and are organized in terms of Reading, Dominance, and Stressed Syllable. These ratio values were used in all the following analyses.

For ease of presentation, mean values of duration, amplitude, and fundamental frequency were also computed for each syllable. To achieve this, first-to-second-syllable ratio values were converted into average values which could then be used for comparison purposes to earlier research. A simple averaging of the acoustic measurements would not result in an appropriate ratio value since an average of the ratio values is not equivalent to a ratio of averaged values. Therefore, to provide more representative values, average ratio values, which were used in the analyses, were converted into first syllable values (x) and second syllable values (y) using the following formulae:

$$x = \frac{\text{first syllable measurement} + \text{second syllable measurement}}{\text{average ratio value} + 1}$$

$$y = (\text{average ratio value})\,(x)$$

These calculated values are presented in Appendix B. For each speaker, values for

duration, amplitude and fundamental frequency are listed for the noun context in Table B1 and for the verb context in Table B2.

The results of the ANOVAs for the ratio data are presented for each speaker and for each parameter. For the purposes of the present investigation, we were particularly interested in any main effect or interaction involving the factor Reading. In the following sections, only significant effects and trends are reported.

4.1 Speaker 1

For Duration, Speaker 1 showed a significant main effect for Stress [$F(1, 12) = 14.86$, $p = .002$], indicating that the first syllable/second syllable duration ratio is significantly greater for words stressed on the first syllable (1.051) compared to words stressed on the second syllable (.496).

For Amplitude, a main effect of Stress [$F(1, 12) = 61.67$, $p = .0001$] indicated that forestressed words have a significantly higher amplitude ratio (1.104) than backstressed words (.937). A significant Stress x Reading interaction [$F(1, 12) = 9.50$, $p = .01$] revealed that forestressed and backstressed words were differentially affected depending on whether they were read as nouns or verbs. Further analyses revealed that forestressed words had higher amplitude ratios in the Noun context (1.132) than in the Verb context (1.076) [$t(7) = 2.31$, $p = .05$] while backstressed words did not have significantly different amplitude ratios in the Noun (.929) compared to the Verb (.945) reading. Finally, a significant Dominance x Reading interaction [$F(1, 12) = 8.24$, $p = .01$] revealed that noun-dominant and verb-dominant words were differentially affected depending on whether they were read as nouns or verbs. Further analyses revealed that noun-dominant words had significantly higher amplitude ratios when read in the Noun list (1.059) compared to the Verb list (1.005) [$t(7) = 2.70$, $p= .03$] while the ratio for verb-dominant words was not significantly different in the Noun list (1.002) compared to the Verb list (1.016).

An analysis of the fundamental frequency data for Speaker 1 revealed no significant effects or interactions.

4.2 Speaker 2

For Duration, Speaker 2 showed a significant main effect for Stress [$F(1, 12) = 14.78$, $p = .002$], indicating that the first syllable/second syllable duration ratio is significantly greater for forestressed words (1.119) compared to backstressed words (.544). In addition, there was a significant trend for Reading [$F(1, 12) = 3.38$, $p = .09$] which

revealed that words read in the Noun context had a slightly higher duration ratio (.855) than words read in the Verb context (.808).

For Amplitude, a main effect of Stress [$F(1, 12) = 206.99$, $p = .0001$] indicated that forestressed words have a significantly higher amplitude ratio (1.177) than backstressed words (.902). In addition, a main effect of Dominance [$F(1, 12) = 5.54$, $p = .04$] showed that noun-dominant words had a higher amplitude ratio (1.062) than verb-dominant words (1.017).

For Fundamental Frequency, a main effect of Stress [$F(1, 12) = 23.08$, $p = .0004$] showed that forestressed words have a significantly higher fundamental frequency ratio (1.149) than backstressed words (.945).

4.3 Speaker 3

For Duration, Speaker 3 showed a significant main effect for Stress [$F(1, 12) = 19.63$, $p = .0008$], indicating that forestressed words have a significantly higher duration ratio (1.146) than backstressed words (.455). In addition, a main effect for Reading [$F(1, 12) = 6.79$, $p = .02$] revealed that words read in the Noun context have a significantly greater duration ratio (.824) than words read in the Verb context (.776).

For Amplitude, a main effect of Stress [$F(1, 12) = 41.42$, $p = .0001$] indicated that forestressed words have a significantly higher amplitude ratio (1.154) than backstressed words (.949).

For Fundamental Frequency, a main effect of Stress [$F(1, 12) = 40.10$, $p = .0001$] indicated that forestressed words have a significantly higher fundamental frequency ratio (1.342) than backstressed words (1.077).

4.4 Speaker 4

For Duration, Speaker 4 showed a significant main effect for Stress [$F(1, 12) = 16.93$, $p = .001$], indicating that the first syllable/second syllable duration ratio is significantly greater for forestressed words (1.309) compared to backstressed words (.532).

For Amplitude, a main effect of Stress [$F(1, 12) = 139.65$, $p = .0001$] indicated that forestressed words have a significantly higher amplitude ratio (1.199) than backstressed words (.896). A significant Dominance x Reading interaction [$F(1, 12) = 5.39$, $p = .04$] revealed that noun-dominant and verb-dominant words were differentially affected depending on whether they were read as nouns or verbs. However, further analyses revealed that none of the individual comparisons were significant: Noun-dominant words read as nouns (1.079) were not significantly different than noun-dominant words read as

verbs (1.054); and verb-dominant words read as nouns (1.014) were not significantly different than verb-dominant words read as verbs (1.044). There was also a trend for the Stress x Reading interaction [$F(1, 12) = 3.22$, $p = .10$]. Further analyses revealed that none of the individual comparisons were significant: forestressed words read as nouns (1.188) were not significantly different compared to forestressed words read as verbs (1.211); and backstressed words read as nouns (.905) were not significantly different compared to backstressed words read as verbs (.886).

For Fundamental Frequency, a main effect of Dominance [$F(1, 12) = 4.95$, $p = .05$] showed that noun-dominant words have a higher fundamental frequency ratio (1.132) than verb-dominant words (1.041). In addition, a main effect of Stress [$F(1, 12) = 4.55$, $p = .05$] indicated that forestressed words unexpectedly have a lower fundamental frequency ratio (1.043) than backstressed words (1.131).

4.5 Speaker 5

For Duration, Speaker 5 showed a significant main effect for Stress [$F(1, 12) = 15.18$, $p = .002$], indicating that the first syllable/second syllable duration ratio is significantly greater for forestressed words (1.236) compared to backstressed words (.559).

For Amplitude, a main effect of Stress [$F(1, 12) = 98.89$, $p = .0001$] indicated that forestressed words have a significantly higher amplitude ratio (1.129) than backstressed words (.943). A significant Stress x Reading interaction [$F(1, 12) = 12.19$, $p = .005$] revealed that forestressed and backstressed words were differentially affected depending on whether they were read as nouns or verbs. Further analyses revealed that forestressed words had significantly higher amplitude ratios when read as nouns (1.159) than when read as verbs (1.100) [$t(7) = 2.81$, $p = .03$], while the ratio for backstressed words was not significantly different in the Noun context (.929) compared to the Verb context (.957).

For Fundamental Frequency, a main effect of Stress [$F(1, 12) = 11.01$, $p = .006$] indicated that forestressed words have a significantly higher fundamental frequency ratio (1.519) than backstressed words (1.042). In addition, there was a trend for a Stress x Reading interaction [$F(1, 12) = 3.63$, $p = .08$]. None of the individual comparisons were significant: ratios for forestressed words were not significantly different in the Noun context (1.668) than in the Verb context (1.370) and ratios for backstressed words were not significantly different in the Noun context (.986) than in the Verb context (1.098).

5. Discussion

The purpose of the present study was to determine whether systematic acoustic differences could be observed in grammatically ambiguous bisyllabic words not exhibiting contrastive stress changes. The present experiment showed slight but consistent acoustic differences for these ambiguous words contingent upon their production as a noun or as a verb. Table 2 summarizes the findings. For all speakers, small, stress-related acoustic differences were found for ambiguous words read in two grammatical categories. The acoustic properties of these ambiguous words seem to depend on their grammatical class usage. In the present study, measurements for each stimulus consisted of a ratio of the first to the second syllable. In such a manner, a valid comparison of the two readings of a stimulus word could be obtained regardless of the idiosyncratic circumstances under which the words were produced.

	Speaker				
	1	**2**	**3**	**4**	**5**
Effect					
Stress	D, A	D, A, F	D, A, F	D, A, F*	D, A, F
Dominance		A		F	
Reading		d	D		
Stress x Reading	A			a*	A, f
Dominance x Reading	A			A	

Table 2. Summary of results for each speaker for all factors considered: Stress (forestressed or backstressed), Dominance (noun dominant or verb dominant), and Reading (noun reading or verb reading). The data for three acoustic parameters are given: Duration (D), Amplitude (A), and Fundamental Frequency (F). Uppercase letters indicate significant effects ($p < .05$) and lowercase letters indicate trends ($p < .10$). An asterisk (*) represents an effect in an unexpected direction.

In general, Stress had a very robust effect for all speakers on all three parameters under investigation. Not surprisingly, the duration, amplitude, and fundamental frequency ratios were greater for forestressed words compared to backstressed words. These results clearly show that a stressed syllable has a longer duration, greater amplitude, and higher fundamental frequency than its unstressed counterpart. Note, however, Speaker 4 does show, for fundamental frequency, a significant effect in the opposite direction. That is, a

stressed syllable for Speaker 4 has a lower fundamental frequency than its unstressed counterpart. Nevertheless, the overall consistent result in the present data is that there is a clear separation between stressed and unstressed syllables in terms of duration, amplitude, and fundamental frequency.

Two speakers showed effects for Dominance: for Speaker 2, the amplitude ratio was greater for noun-dominant than for verb-dominant words; for Speaker 4, the fundamental frequency ratio was greater for noun-dominant than for verb-dominant words. Regardless of stress placement, word stimuli that were more frequent as nouns in English (e.g., *poison, debate*) showed a significantly different stress pattern than word stimuli that were more frequent as verbs (e.g., *notice, escape*). Speakers 2 and 4 acknowledged the frequency dominance of ambiguous words by increasing amplitude or fundamental frequency cues in the first syllable for noun-dominant words and in the second syllable for verb-dominant words. These effects suggest that these speakers maximized the difference between noun- and verb-dominant words in conformity with the lexical distribution of English in which the majority of bisyllabic nouns are stressed on the first syllable and the majority of bisyllabic verbs on the second syllable.

For the present study, we were particularly interested in a main effect or interaction involving the factor Reading. In the present experiment, each stimulus was read both as a noun and a verb. Table 2 shows that every speaker in this study displayed such an effect of Reading. The data clearly demonstrate significant acoustic differences between the noun reading and the verb reading of the stimuli. The main effect for the factor Reading involved durational cues. Speakers 2 and 3 produced stimulus tokens with higher duration ratios when spoken as nouns compared to verbs. That is, the same words were produced with higher first-to-second syllable duration ratios when produced as nouns compared to lower first-to-second syllable duration ratios when produced as verbs. Thus, durational information is systematically exploited by at least some speakers in the production of grammatically ambiguous lexical items. Results of previous studies investigating the acoustic correlates of lexical stress in English found that variations in duration as well as fundamental frequency, amplitude, and formant structure contribute to the perception of lexical stress (e.g., Gay, 1978; Lehiste, 1970). Although it has been found that stressed syllables tend to have longer durations than unstressed syllables, fundamental frequency has typically been shown to be a more effective cue to the contrastive stress patterns found in English (e.g., Fry, 1958; 1965; Lieberman, 1960; 1967). It has also been noted, however, that intensity information in combination with durational cues may be a sufficient cue in the production and perception of stress in English (e.g., Beckman, 1986).

Nevertheless, the results of the two speakers in the present experiment who systematically vary duration as a function of reading suggest that durational cues may be more robust in signaling grammatical class differences.

Three other speakers showed a number of significant interactions involving the Reading factor. Speakers 1 and 5 showed Reading by Stressed Syllable interactions while Speakers 1 and 4 showed Reading by Dominance interactions. Most of the significant interactions involved amplitude cues.

Speakers 1 and 5 adjusted their productions as a function of Reading and Stressed Syllable. Forestressed and backstressed words were differentially affected depending on their grammatical function. This interaction, however, affected forestressed words more than backstressed words. Forestressed words had significantly greater amplitudinal ratios in the noun compared to the verb reading while backstressed words showed no significant differences as a result of reading. The grammatical class of a lexical item seems to affect amplitudinal information in forestressed words more than backstressed words.

Speakers 1 and 4 adjusted their productions as a function of Reading and Dominance. Noun and verb dominant words were differentially affected depending on their grammatical function. The frequency dominance of the grammatical categories of a lexical item affected its production in terms of amplitudinal measures. This interaction affected noun-dominant words more than verb-dominant words. Noun-dominant words had significantly greater amplitude ratios in the Noun reading compared to the verb reading while verb-dominant words showed no significant effects as a result of reading.

First, it should be noted that the existence of such interactions suggests that speakers are sensitive not only to lexical stress but also to the individual frequency of occurrence of the separate grammatical categories of a lexical item. This latter observation that noun dominant words are affected differently than verb dominant words would support an organization of the lexicon in which there is a separate tabulation of frequency information by grammatical category. Additional experimentation is needed to further substantiate such claims.

Second, the interactions involving Reading generally show significant differences for forestressed compared to backstressed words and noun dominant compared to verb dominant words. It has been suggested that nouns are unmarked relative to verbs (Greenberg, 1978; Gentner, 1978). If this is the case, than noun dominant words and words with typical noun stress placement characteristics may also be unmarked. Since verbs and backstressed words are marked, it may be more salient for speakers to produce contrasts involving unmarked words. Previous research examining vowel differences

between nouns and verbs also showed more sizable effects in noun compared to verb stimuli (Sereno and Jongman, 1990).

Third, the interactions observed in the present study generally involved amplitudinal cues. Although a complex of acoustic features involving duration, amplitude, fundamental frequency, and spectral reduction have been shown to be effective cues in the perception and production of stress in English, many investigations have demonstrated that fundamental frequency patterns may be sufficient cues. It may be surprising, therefore, that in the present experiment there were no significant comparisons between noun and verb readings involving fundamental frequency measures. It is possible that fundamental frequency cues are reserved to signal lexical stress information whereas duration and amplitude cues are used to signal grammatical category membership.

It should also be noted that Speaker 4's productions exhibit two effects that are contrary to expectations. Both of these effects are trends and do not reach significance. First, Speaker 4 has a tendency to produce first-syllable stressed words with slightly lower fundamental frequency ratios than second-syllable stressed words. Second, for amplitude measurements, although Speaker 4 showed a trend for the reading by stressed syllable interaction, neither of the individual comparisons (noun reading versus verb reading of forestressed words; noun reading versus verb reading of backstressed words) were significant. Earlier studies have occasionally mentioned effects in the unexpected direction (e.g., Morton and Jassem, 1965; Beckman, 1986). However, for the cues to be perceptually salient, these effects usually involved quite substantial changes in the acoustic parameters. The present results do not show such dramatic changes.

Taken together, the present findings demonstrate that systematic, significant differences in speakers' production of syntactically ambiguous words. Words read as nouns emphasize first-syllable stress and words read as verbs emphasize second-syllable stress. These production data demonstrate that the acoustic correlates of lexical stress initially noted in words exhibiting contrastive stress are retained as traces in bisyllabic grammatically ambiguous words in English. In the present study, duration and amplitude appear to be most robust cues for signaling differences in grammatical function which is in contrast to lexical stress cues emphasizing fundamental frequency information.

The differences observed in this study are quite consistent. All of the significant effects involving the factor Reading go in the direction predicted by the structural characteristics of the language. English manifests a typical pattern of forestressed nouns and backstressed verbs. The present experiment shows that these patterns may be important in differentiating between different grammatical classes. These results suggest that speakers

are not only aware of the relation between grammatical category and stress placement in English but also seem to use this information in the production of ambiguous lexical items.

The differences found in the present study cannot be attributed to the sentential position of nouns and verbs. Sorensen, Cooper, and Paccia (1978) demonstrated that many durational differences between nouns and verbs can be accounted for in terms of differences in position within a constituent. Due to relatively strict word order constraints in English, perhaps systematic acoustic differences could have arisen at or at least remained consistent with sentential level phonological information (see, for example, Kelly, 1992; Sorensen et al., 1978). However, the present results were found for stimuli read from a nouns-only or verbs-only list, with identical 'contextual' influences. The present results, therefore, suggest that grammatical category membership itself, independent of sentential information, can have a significant and systematic effect on the modulation of stress in English. Perhaps these data provide a possible clue to grammatical category assignment in English.

A crucial question, of course, is whether listeners are sensitive to these subtle acoustic cues to grammatical category. Further research is planned which will examine whether subjects use these acoustic cues in recognition. A possible extension of the present study would be to investigate whether listeners can accurately determine the grammatical class of an ambiguous word produced as either a noun or a verb.

In conclusion, the present results begin to explore the way in which grammatical category affects the acoustic characteristics of words. The particular issue investigated was the effect of grammatical category on lexical stress cues such as duration, amplitude, and fundamental frequency. The present study suggest that there are some systematic acoustic differences between a single stimulus item read as a noun, on the one hand, and a verb, on the other. These data, derived from an analysis of speech production, seem to be in accord with well-established acoustic cues for contrastive stress which have shown to be salient in speech perception. An obvious next step is to determine whether such differences are functional for the listener.

References

Beckman, M. (1986). *Stress and Non-Stress Accent.* Dordrecht: Foris.

Bolinger, D. (1958). A theory of pitch accent in English. *Word* 14, 109-149.

Chomsky, N. and Halle, M. (1968). *The Sound Pattern of English.* New York: Harper and Row.

Cutler, A. (1991). Exploiting prosodic possibilities in speech segmentation. In G.

Altmann (ed.), *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives.* Cambridge, MA: MIT Press.

Cutler, A., and Clifton, C. (1984). The use of prosodic information in word recognition. In H. Bouma and D.G. Bouwhuis (eds.), *Attention and Performance X: Control of language processes.* Hillsdale: Lawrence Erlbaum, pp. 183-196.

Francis, W.N., and Kucera, H. (1982). *Frequency Analysis of English Usage: Lexicon and Grammar.* Boston, Mass.: Houghton Mifflin Co.

Fry, D.B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America,* 27, 765-768.

Fry, D.B. (1958). Experiments in the perception of stress. *Language and Speech* , 1, 126-152.

Fry, D.B. (1965). The dependence of stress judgments on vowel formant structure. In E. Zwirner and W. Bethge (eds.), *Proceedings of the 6th International Congress of Phonetic Sciences.* Basel: Karger.

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

Gentner, D. (1978). On relational meaning: The acquisition of verb meaning. *Child Development,* 49, 988-998.

Greenberg, J. (1978). *Universals of Human Language.* Stanford:Stanford University Press.

Hayes, B. (1981). *A Metrical Theory of Stress Rules.* Doctoral Dissertation, Cambridge, MA: MIT Press.

Kawamoto, A., A.H., Farrar, W.T., and Overbeek, M. (1993). Effects of grammatical category and stress pattern on on-line processing of bisyllabic words, (manuscript).

Kelly, M.H. (1988). Phonological biases in grammatical category shifts. *Journal of Memory and Language* , 27, 343-358.

Kelly, M. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review,* 99, 349-364.

Kelly, M.H., and Bock, J.K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance* , 14, 389-403.

Lehiste, I. (1970). *Suprasegmentals.* Cambridge, Mass: MIT Press.

Liberman, M.Y., and Prince, A.S. (1977). On stress and linguistic rhythm. *Linguistic Inquiry* , 8, 249-336.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America* , 32, 451-454.

Lieberman, P. (1967). *Intonation, Perception, and Language*. Cambridge, Mass.: MIT Press.

Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* , 35, 1773-1781.

Morton, J. and Jassem, W. (1965). Acoustic correlates of stress. *Language and Speech* , 8, 159-181.

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

Pierrehumbert, J. (1980). *The Phonology and Phonetics of English Intonation*. Bloomington: Indiana University Linguistics Club.

Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. (1986). *A comprehensive grammar of the English language*. London: Longman.

Reetz, H. (1989). A fast expert program for pitch extraction. *Proceedings of Eurospeech: Paris 1989*, 1, 476-479.

Sereno, J.A. (1986). Stress pattern differentiation of form class in English. *Journal of the Acoustical Society of America* , 79, S36.

Sereno, J.A. and Jongman, A. (1990). Phonological and form class relations in the lexicon. *Journal of Psycholinguistic Research*, 19(6), 387-404.

Sorensen, J., Cooper, W., and Paccia, J. (1978). Speech timing of grammatical categories. *Cognition*, 6, 135-153.

Sherman, D. (1975). Noun-verb stress alternation: An example of lexical diffusion of sound change in English. *Linguistics* , 159, 43-71.

Treiman, R. (1989). The internal structure of the syllable. In G. Carlson and M. Tanenhaus (eds.), *Linguistic Structure in Language Processing*. Dordrecht: Kluwer.

Webster's Seventh New Collegiate Dictionary (1963). Springfield, MA: G. and C. Merriam Company.

7. Appendix A

	Noun reading				Verb reading			
	Noun dominant		Verb dominant		Noun dominant		Verb dominant	
Speaker	Fore	Back	Fore	Back	Fore	Back	Fore	Back
1	.972	.493	1.072	.515	1.16	.525	.997	.453
2	1.155	.635	1.153	.478	1.08	.567	1.087	.498
3	1.29	.482	1.055	.47	1.188	.43	1.05	.438
4	1.28	.577	1.317	.528	1.388	.56	1.25	.463
5	1.413	.538	1.272	.575	1.165	.57	1.093	.555

Table A1. Mean first/second syllable duration ratios for each speaker. Ratios are organized in terms of Reading (noun reading, verb reading), Dominance (noun dominance, verb dominance) and Stress Syllable (Forestressed, Backstressed).

	Noun reading				Verb reading			
	Noun dominant		Verb dominant		Noun dominant		Verb dominant	
Speaker	Fore	Back	Fore	Back	Fore	Back	Fore	Back
1	1.158	.96	1.107	1.085	1.068	.943	1.085	.947
2	1.19	.92	1.157	.885	1.212	.925	1.147	.877
3	1.17	.943	1.12	.965	1.21	.942	1.115	.945
4	1.23	.928	1.145	.882	1.21	.897	1.212	.875
5	1.168	.933	1.15	.925	1.13	.94	1.07	.975

Table A2. Mean first/second syllable amplitude ratios for each speaker. Ratios are organized in terms of Reading (noun reading, verb reading), Dominance (noun dominance, verb dominance) and Stress Syllable (Forestressed, Backstressed).

	Noun reading				Verb reading			
	Noun dominant		Verb dominant		Noun dominant		Verb dominant	
Speaker	Fore	Back	Fore	Back	Fore	Back	Fore	Back
1	.847	.855	.812	.858	.825	.87	.842	.845
2	1.21	.98	1.058	.967	1.17	.928	1.16	.905
3	1.34	1.075	1.265	1.097	1.408	1.097	1.355	1.04
4	1.043	1.225	1.008	1.093	1.092	1.17	1.028	1.035
5	1.783	.925	1.553	1.048	1.538	.975	1.202	1.22

Table A3. Mean first/second syllable fundamental frequency ratios for each speaker. Ratios are organized in terms of Reading (noun reading, verb reading), Dominance (noun dominance, verb dominance) and Stress Syllable (Forestressed, Backstressed).

8. Appendix B

Noun Reading

	Speaker	Duration (ms)		Amplitude (dB)		F0 (Hz)	
		1st	2nd	1st	2nd	1st	2nd
	1	326	335	45.4	39.2	113	133
	2	316	273	37.6	31.7	121	100
Noun dominant	3	341	265	45.6	39.0	239	179
forestressed	4	392	306	44.3	36.0	185	178
	5	376	266	43.7	37.4	243	136
	1	216	438	41.6	43.4	106	124
	2	248	390	34.9	38.0	110	113
Noun dominant	3	182	378	41.3	43.9	215	201
backstressed	4	230	399	37.0	39.8	205	168
	5	215	400	38.7	41.5	186	201
	1	340	317	44.4	40.0	108	133
	2	319	277	37.4	32.4	120	113
Verb dominant	3	303	287	44.3	39.6	224	177
forestressed	4	395	300	41.4	36.2	184	183
	5	350	275	44.2	38.4	232	150
	1	254	492	38.7	43.1	101	118
	2	218	456	33.1	37.4	107	110
Verb dominant	3	225	479	41.7	43.2	216	197
backstressed	4	251	475	37.2	42.3	198	181
	5	262	456	39.4	42.5	195	187

Table B1. For each speaker, mean values for duration (ms), amplitude (dB), and fundamental frequency (Hz) are presented for each syllable. Measurements are organized in terms of Dominance (Noun dominant and Verb dominant) and Stress (forestressed and backstressed).

Verb Reading

	Speaker	Duration (ms)		Amplitude (dB)		F0 (Hz)	
		1st	2nd	1st	2nd	1st	2nd
Noun dominant forestressed	1	350	302	45.7	42.8	107	130
	2	326	302	37.8	31.2	126	108
	3	339	285	47.1	39.0	248	176
	4	401	289	43.3	35.9	187	171
	5	324	278	44.5	39.3	240	156
Noun dominant backstressed	1	219	417	43.3	45.9	97	113
	2	239	422	35.2	38.1	112	122
	3	173	402	41.6	44.1	212	193
	4	231	412	35.6	39.6	203	173
	5	208	364	39.1	41.6	189	194
Verb dominant forestressed	1	321	322	47.2	43.7	107	127
	2	319	294	37.6	32.9	122	105
	3	287	273	45.5	40.8	233	172
	4	364	291	43.4	35.8	195	191
	5	327	299	42.6	39.8	201	167
Verb dominance backstressed	1	219	484	43.6	46.0	92	109
	2	242	485	33.0	37.7	110	121
	3	197	449	41.4	43.7	207	199
	4	223	482	36.5	41.7	194	187
	5	238	429	40.8	41.8	204	167

Table B2. For each speaker, mean values for duration (ms), amplitude (dB), and fundamental frequency (Hz) are presented for each syllable. Measurements are organized in terms of Dominance (Noun dominant and Verb dominant) and Stress (forestressed and backstressed).