*Acoustic Correlates of Grammatical Class**

JOAN A. SERENO
ALLARD JONGMAN
Cornell University

KEY WORDS

lexical stress

acoustic cues

grammatical class

noun

verb

ABSTRACT

The present experiment investigates acoustic correlates of grammatical class in English. Results of previous studies examining the acoustic correlates of stress have established that variations in duration, amplitude and fundamental frequency are cues to lexical stress. The present study investigates whether systematic acoustic differences would also be observed in words which maintain a constant stress pattern across syntactic function. In the present experiment, five speakers were recorded producing 16 grammatically ambiguous lexical items such as *answer* or *design*. Stimuli were read either as a noun or a verb. Measurements for each stimulus consisted of a ratio of the first to the second syllable for the three acoustic parameters: duration, amplitude, and fundamental frequency. The data show slight but consistent acoustic differences in speakers' production of syntactically ambiguous words contingent upon their production as a noun or as a verb. Duration and amplitude cues appear to be the most robust cues for signaling these differences in grammatical function. These results are in accord with recent evidence documenting the paradigmatic stress pattern for grammatical class in English.

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 & Halle, 1968; Liberman & Prince, 1977; Hayes, 1981) and sentential levels (Pierrehumbert, 1980; Beckman, 1986). 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

* We would like to thank the Brown Phonetics Laboratory for the use of its recording facilities and Annemie Witjes, Cisca Custers, Toine Thissen, and Sylvain Laroche for assistance. Technical help by John Mertus and Henning Reetz is gratefully acknowledged. Finally, comments by Shari Speer, an anonymous reviewer, and the editors were helpful. A preliminary report of this research appeared as Sereno and Jongman (1993).

Address all correspondence to Joan A. Sereno, Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853 USA. E-mail: joan.sereno@cornell.edu

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 1950s, 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 Lehiste, 1970; Gay, 1978). 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 have been shown to contribute to the perception of stress (Fry, 1955, 1958; Bolinger, 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 (F_0) appears to be the most predominant cue to perceived stress, variations in duration, amplitude and formant structure also systematically contribute to stress judgments.

The relative importance of each of these parameters, however, varies with the position of the lexical item in the sentence, suggesting a number of interactions (Morton & Jassem, 1965; Gay, 1978; Nakatani & Aston, 1978; Beckman, 1986). It seems that, in nuclear position, F_0 does show up as a cue, with amplitude and duration as secondary cues. However, in non-nuclear position, amplitude and duration cues become more salient, often surpassing fundamental frequency as a cue. Beckman (1986) therefore suggests that fundamental frequency may be a primary cue to accent rather than stress. It is not surprising that the relative effectiveness of fundamental frequency depends on the position of the word in the sentence.

The present study will examine acoustic properties of isolated words bearing nuclear accent. In the production of isolated words, a complex of acoustic cues, including duration, intensity, fundamental frequency, and spectral composition, appears to collectively contribute to the perception of stress.

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 (trochaic) and verbs are backstressed (iambic) is quite pervasive in English. Although the linguistic literature commonly references this fact (e.g., Chomsky & Halle, 1968; Sherman, 1975; Liberman & Prince, 1977), few attempts were made to quantify the pattern statistically.

Recently, however, a number of studies have systematically investigated this striking noun-verb stress asymmetry (Cutler & Clifton, 1984; Sereno, 1986; Kelly, 1988; Kelly & Bock, 1988; Sereno & Jongman, 1993). In a lexical analysis of English, Sereno

TABLE 1

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

		Number of syllables							
Type	*n*	1	2	3	4	5	6	7	8
Nouns	3,858	536	1,425	1,061	586	210	35	3	2
		(14)	(37)	(28)	(15)	(5)	(1)	(0)	(0)
Verbs	999	187	523	222	62	5			
		(19)	(52)	(22)	(6)	(1)			
Noun/Verb Ambiguous	1694	1,014	570	97	13				
		(60)	(34)	(6)	(1)				

(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 & Kucera, 1982) in terms of number of syllables for pure nouns (words used only as nouns in the Brown Corpus), pure verbs (words used only as verbs in the Brown Corpus), and ambiguous items (words with both noun and verb instances in the Brown Corpus). Lexical items included words with a frequency of five per million or greater. For each item, the number of syllables was established by its transcription in Webster's Seventh New Collegiate Dictionary (1963).

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 bisyllabic 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. Words are categorized as either 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.

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

TABLE 2

Number of bisyllabic noun/verb ambiguous words in each combination of stress placement and grammatical category dominance in Francis and Kucera (1982)

	Dominant Class			
Stress Location	*Noun*	*Verb*	*Equi*	*Total*
Forestress	293	85	14	392
Backstress	66	51	2	119
Variable	31	27	1	59
				570

from the first to the second syllable (for a list of words, see Quirk, Greenbaum, Leech, & Svartvik, 1986, pp. 1566 – 1567). 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 words do not shift stress placement with grammatical class membership. Nevertheless, they show consistent stress placement depending on their dominant grammatical class usage. That is, for these ambiguous stimuli, there is a significant association between stress placement and dominance, $\chi^2 (1, N = 485) = 19.97$, $p < .001$. Forestressed stimuli are much more likely to be noun-dominant (75%) than verb-dominant (22%) while backstressed stimuli are only marginally more noun-dominant (55%) than verb-dominant (43%).

Production studies confirmed the influence of paradigmatic stress placement in English. Recent findings by Kelly and colleagues (see Kelly, 1992, for a review) show that subjects asked to read aloud sentences containing bisyllabic nonwords produced these nonwords more often with first-syllable stress when the syntactic context indicated the nonword functioned as a noun, and with second-syllable stress when the context indicated a verb (Kelly & 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 study by Kawamoto, Farrar, and Overbeek (1993). Kawamoto et al. (1993) found that naming latencies to bisyllabic words exhibited a systematic effect when preceded by an appropriate syntactic context. They showed that stimuli with stress patterns that were consistent with that predicted by their grammatical category (i.e., first-syllable stress for nouns and second-syllable stress for verbs) were named more quickly. 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 at least that 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 (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 would also be observed in these words which do not exhibit stress changes. The presence of such differences would suggest that speakers are not only aware of the relation between grammatical category and stress placement but also seem to use this information in the production of ambiguous lexical items. 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 stress-related acoustic differences in grammatically ambiguous words read in two different grammatical class contexts may be detectable.

METHOD

Subjects

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

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 occur as both nouns or verbs in English, without change in lexical stress. None of the stimuli have any occurrences as other grammatical categories, except for the word 'welcome' which also occurs as an adjective and exclamation[1] (Francis & Kucera, 1982). Stimulus items were balanced both with respect to form class frequency as well as stress

[1] 'Welcome' is the only stimulus with occurrences other than noun or verb (Francis & Kucera, 1982). In the Brown corpus, 'welcome' occurred an additional 22 times per million as an adjective and exclamation. The percentages for 'welcome' in the table represent only its noun and verb usages.

TABLE 3

The 16 stimulus words are listed with their frequency of occurrence (per million) as nouns and verbs (Francis & Kucera, 1982). Percentages represent proportion of noun and verb frequency of total frequency of occurrence

Stimulus Class	***Class***	
	Noun	*Verb*
NOUN-DOMINANT		
FIRST-SYLLABLE STRESS		
favor	63	49
poison	11	7
practice	132	40
struggle	57	36
$\bar{X}$	66 (67%)	33 (33%)
SECOND-SYLLABLE STRESS		
control	220	95
debate	36	10
dispute	37	14
report	205	184
$\bar{X}$	125 (63%)	76 (38%)
VERB-DOMINANT		
FIRST-SYLLABLE STRESS		
handle	22	81
notice	39	84
rescue	9	14
welcome	13	33
$\bar{X}$	21 (28%)	53 (72%)
SECOND-SYLLABLE STRESS		
embrace	6	18
escape	24	69
neglect	8	28
reply	35	75
$\bar{X}$	18 (27%)	48 (73%)

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*); noun-dominant and backstressed (e.g., *debate*); verb-dominant and forestressed (e.g., *notice*); and verb-dominant and backstressed (e.g., *escape*). The noun-dominant and forestressed stimuli had an average noun usage frequency of 66 per million and an average verb frequency usage of 33 per million. The noun-dominant and backstressed stimuli had an average noun usage frequency of 125 per million and an average verb frequency usage of 76 per million. The verb-dominant and forestressed stimuli had an average noun usage frequency of 21 per million and an average verb frequency usage of 53 per million. And the verb-dominant and backstressed stimuli had an average noun usage frequency of 18 per million and an average verb frequency usage of 48 per million.

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 pure nouns and the verb context consisted of a list of 75 pure verbs. The noun context were words occurring only as nouns in the Brown Corpus, with no occurrences as other grammatical categories (Francis & Kucera, 1982). These pure nouns included words such as 'balcony' and 'soprano' The verb context were words occurring only as verbs in the Brown Corpus, with no occurrences as other grammatical categories (Francis & Kucera, 1982). These pure verbs included words such as 'circulate' and 'prohibit' The stimuli selected for the noun context and verb context lists were carefully controlled for stress placement in order to prevent biasing due to list composition. Both the noun context and verb context lists consisted of five monosyllabic words, 60 words with three syllables, nine words with four syllables, and one word with five syllables. For the 70 multisyllabic words in each list, 34 had first-syllable stress, 34 had second-syllable stress, and two had third-syllable stress.

The 16 stimuli were 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 subjects were told they would be reading either "a list of nouns in English" (noun context) or "a list of verbs in English" (verb context). The order of the contexts in which the stimulus words were read was approximately 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.

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 waveforms 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. Waveforms and spectrograms of 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 (*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 lists of 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 also 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 (F_0) contours were obtained for each syllable, using a pitch extraction algorithm based on an autocorrelation procedure (Reetz, 1989). Since F_0 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 F_0) per syllable was used. A ratio of first-to-second syllable fundamental frequency was then computed for each item.

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.

The raw data are presented in Appendix B. For each speaker, mean values for duration, amplitude and fundamental frequency are listed for the noun and verb contexts.

The ratio data were analyzed for all subjects as well as for each individual subject. Separate analyses were performed for each acoustic parameter: duration, amplitude, and fundamental frequency. The overall analysis will be presented first, followed by the analysis of each subject's data. 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.

Overall Analysis

A three-way (Stress × Dominance × Reading) repeated-measures ANOVA, with Reading as the within variable, was conducted for all subjects' ratio data. Items were treated as a random variable. Separate ANOVAs were conducted for each acoustic parameter: amplitude, duration, and fundamental frequency.

For Duration, there was a significant main effect for Stress, $F(1,76) = 98.79$, $p = .0001$, indicating that the first syllable/second syllable duration ratio is significantly greater for words stressed on the first syllable (1.1896) than for words stressed on the second syllable (0.5260).

For Amplitude, there was a significant main effect for Stress, $F(1,76) = 230.82$, $p = .0001$, indicating that the first syllable/second syllable amplitude ratio is significantly greater for words stressed on the first syllable (1.1667) than for words stressed on the second syllable (0.9276). There was also a significant main effect for Dominance, $F(1,76) = 4.49$, $p = .037$, indicating that noun-dominant words had a significantly higher amplitude ratio (1.0638) than verb-dominant words (1.0305). In addition, there was a significant Stress × Reading interaction, $F(1,76) = 3.89$, $p = .05$, indicating that forestressed and backstressed words were differentially affected depending on whether they were read as nouns or verbs. Further analyses revealed that forestressed words tended to have higher amplitude ratios in the Noun context (1.1778) than in the Verb context (1.1555), $t(39) = 1.59$, $p = .06$ while backstressed words did not have significantly different amplitude ratios in the Noun (.9237) and Verb (.9315) readings.

For Fundamental Frequency, there was a significant main effect for Stress $F(1,76) = 11.09$, $p = .001$, indicating that the first syllable/second syllable fundamental frequency ratio is significantly greater for words stressed on the first syllable (1.2078) than for words stressed on the second syllable (1.0341).

Separate analyses for each subject are presented below.

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) than for words stressed on the second syllable (0.496).

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 (0.937). A significant Stress × 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) and Verb (.945) readings. Finally, a significant Dominance × 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 = .031, while the ratio for verb-dominant words was not significantly different in the Noun (1.002) and Verb (1.016) lists.

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

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) than for backstressed words (0.544). In addition, the Reading effect approached significance, F (1,12) = 3.38, p = .09 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 (0.902). 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 (0.945).

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 (0.455). 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 (0.949).

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).

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) than for backstressed words (0.532).

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 (0.896). A significant Dominance × 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 × Reading interaction, F (1,12) = 3.22, p = .10

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).

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) than for backstressed words (0.559).

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 (0.943). A significant Stress × 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 = .026, 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 × 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 (0.986) and Verb context (1.098).

DISCUSSION

The purpose of the present study was to determine whether systematic acoustic differences could be observed in grammatically ambiguous bisyllabic words not

TABLE 4

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

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

exhibiting phonemic stress changes. For all speakers, there were small, stress-related acoustic differences found between noun and verb readings of ambiguous words. The acoustic properties of these ambiguous words seem to depend on their grammatical class usage. Table 4 summarizes the findings. In the present study, measurements for each stimulus consisted of a ratio of the first to the second syllable for the three acoustic parameters: duration, amplitude, and fundamental frequency. 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.

In general, Stress had a very robust effect on all three parameters under investigation. This was true for the overall analysis as well as the separate analyses for each of the speakers. Not surprisingly, the duration, amplitude, and fundamental frequency ratios were greater for forestressed words than for 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, that 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.

The overall analysis also showed a significant main effect for Dominance. For amplitude measures, noun-dominant stimuli had higher amplitude ratios than verb-dominant stimuli. For all speakers, stimuli that were more frequent as nouns in English (e.g., *poison*, *debate*) showed significantly different amplitude ratios than word stimuli that were more frequent as verbs (e.g., *notice*, *escape*). Moreover, two individual speakers

showed significant 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. Speakers 2 and 4 reacted to 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. The significant Dominance effects suggest that 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 4 shows that every speaker in this study displayed a significant effect or trend for Reading. The data demonstrate notable acoustic differences between the noun reading and verb reading of the stimuli. The main effects for the factor Reading involved durational information. Speaker 3 produced stimulus tokens with significantly higher duration ratios when spoken as nouns compared to verbs while Speaker 2 showed a trend in the same direction. 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 seems to be systematically exploited by at least some speakers in the production of grammatically ambiguous lexical items.

Although, in nuclear position, stressed syllables tend to have higher fundamental frequency than unstressed syllables (e.g., Fry, 1958; 1965; Lieberman, 1960; 1967), variations in duration also contribute to the perception of lexical stress (e.g., Lehiste, 1970; Gay, 1978). Moreover, it has been noted that duration information in combination with intensity cues may be a sufficient cue in the production and perception of stress in English (e.g., Beckman, 1986). The present results showing significant variations in duration as a function of reading suggest that durational cues may be robust in signaling grammatical class differences.

The overall analysis additionally showed a significant interaction involving the Reading factor. For amplitude measures, speakers showed a significant Reading by Stressed Syllable interaction. In addition to this overall interaction, Speakers 1 and 5 individually showed significant Reading by Stressed Syllable interactions for amplitude measures. The present results show that speakers' productions systematically varied as a function of the stressed syllable, with forestressed and backstressed words differentially affected depending on their grammatical function. Subsequent comparisons revealed that these interactions 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 amplitude in forestressed words more than backstressed words.

Two speakers (Speakers 1 and 4) also showed a significant Reading by Dominance interaction for amplitude measures. 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 the relative amplitude of its syllables.

Again, this interaction affected noun-dominant words more than verb-dominant words. Subsequent comparisons of the noun and verb readings revealed a significant difference for noun-dominant stimuli but not for verb-dominant stimuli. Noun-dominant stimuli showed significantly higher amplitude ratios for stimuli in the noun reading compared to the verb reading while verb-dominant stimuli showed no differences between the readings.

First, it should be noted that the existence of such interactions involving Stressed Syllable and Dominance 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 may 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 only for forestressed (contrasted to backstressed) and noun-dominant (contrasted to verb-dominant) words. It should be noted that in the lexical analysis of bisyllabic English words (Table 2), forestressed words (compared to backstressed words) and noun-dominant words (compared to verb-dominant words) both show a more skewed distribution. Also a majority of the grammatically ambiguous words are forestressed and noun-dominant (51%). Moreover, in the present study, the noun-dominant stimuli had an average noun frequency of 95 per million and an average verb frequency of 54 per million. Although the verb-dominant stimuli had a similar, but reversed, proportion of noun (20 per million) and verb frequency (50 per million), these verb-dominant stimuli had much lower total stimulus frequency (noun-dominant: 149 per million; verb-dominant: 70 per million). These observations suggest a more foundational role for forestressed, noun-dominant stimuli. It has also been suggested that nouns are unmarked relative to verbs (Greenberg, 1978; Gentner, 1978). If this is the case, then noun-dominant words and words with typical noun stress placement characteristics may also be unmarked. Since verbs (relative to nouns) and backstressed words (relative to forestressed 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 & Jongman, 1990).

Third, the interactions observed in the present study generally involved amplitude. Although many investigations have demonstrated that fundamental frequency patterns may be sufficient cues in nuclear position to the perception and production of stress in English, there were no significant comparisons between noun and verb readings involving fundamental frequency in the present study. Beckman and colleagues (e.g., Beckman, 1986; Beckman & Pierrehumbert, 1986; Beckman & Edwards, 1994) have argued that in English fundamental frequency may not signal stress but is really a cue to accent. Duration, amplitude, and vowel quality may instead be the primary cues to stress. Nevertheless, in nuclear position where stress and accent are associated with the same syllable, fundamental frequency often contributes to the perception of stress. However, the present study, in which isolated words were used, shows that even in nuclear position, the differences between nouns and verbs that are most often present involve duration and amplitude, not fundamental frequency. It may be the case then that in nuclear position duration and amplitude are used to signal grammatical category membership.

It should also be of interest that Speaker 4's productions exhibit two effects that are contrary to expectations. First, Speaker 4 produced first-syllable stressed words with *lower* fundamental frequency ratios than second-syllable stressed words. Second, for amplitude measurements, Speaker 4 showed a trend for the Reading x Stressed Syllable interaction. Earlier studies have occasionally mentioned effects in the unexpected direction (e.g., Morton & 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 slight but consistent acoustic differences in speakers' production of syntactically ambiguous words contingent upon their production as a noun or as a verb. Words read as nouns favor first-syllable stress and words read as verbs favor second-syllable stress. These production data demonstrate that the acoustic correlates of lexical stress initially noted in words exhibiting phonemic stress contrasts are retained as traces in bisyllabic grammatically ambiguous words in English. In the present study, duration and amplitude appear to be the most robust cues for signaling differences in grammatical function.

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 English which 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.

It has been demonstrated that many durational differences between nouns and verbs can be accounted for in terms of differences in position within a constituent (Sorensen, Cooper, & Paccia, 1978). Due to relatively strict word order constraints in English, it is possible that systematic acoustic differences could have originated at or at least remained consistent with sentential level phonological information (see, for example, Sorensen et al., 1978; Kelly, 1992). However, the present results were found for stimuli read from a nouns-only or verbs-only list, with identical 'contextual' influences. The acoustic differences found in the present study cannot be attributed to the sentential position of nouns and verbs. 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.

The use of stress cues to signal grammatical category is relevant to aspects of language acquisition. Recent research (Gleitman, 1990; Hirsh-Pasek, Gleitman, Gleitman, Golinkoff, Naigles, 1988; Naigles, 1990; Soja, 1992; Bloom, 1994) indicates that information about grammatical class may facilitate word learning, a process called "syntactic bootstrapping" These studies suggest that children as young as two years of age are sensitive to grammatical class differences which seem to aid them in acquiring word meaning.

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 suggests 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 stress which have been shown to be salient in speech perception. An obvious next step is to determine whether such differences are functional for the listener.

(Received: July 19, 1994; accepted February 26, 1995)

REFERENCES

BECKMAN, M. (1986). *Stress and non-stress accent.* Dordrecht: Foris.

BECKMAN, M.E., & EDWARDS, J. (1994). Articulatory evidence for differentiating stress categories. In P.A. Keating (Ed.), *Phonological structure and phonetic form: Papers in Laboratory Phonology III* (pp. 7 – 34). Cambridge, UK: Cambridge University Press.

BECKMAN, M.E. & PIERREHUMBERT, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, **3**, 255 – 309.

BLOOM, P (1994). Possible names: The role of syntax-semantics mappings in the acquisition of nominals. *Lingua*, **92**, 1 – 33.

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

CHOMSKY, N. & 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* (pp. 105 – 122). Cambridge, MA: MIT Press.

CUTLER, A., & CLIFTON, C. (1984). The use of prosodic information in word recognition. In H. Bouma & D.G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 183 – 196). Hillsdale: Lawrence Erlbaum

FRANCIS, W.N., & 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 & W. Bethge (Eds.), *Proceedings of the 6th International Congress of Phonetic Sciences* (pp. 306 – 311). 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.

GLEITMAN, L.R. (1990). The structural sources of word meaning. *Language Acquisition*, **1**, 3 – 55.

GREENBERG, J. (1978). *Universals of human language.* Stanford: Stanford University Press.

HAYES, B. (1981). *A metrical theory of stress rules.* Cambridge, Mass.: MIT Press.

HIRSH-PASEK, K., GLEITMAN, H., GLEITMAN, L.R., GOLINKOFF, R., & NAIGLES, L. (1988). *Syntactic bootstrapping: Evidence from comprehension.* Paper presented at the Boston University Conference on Language Development, Boston.

KAWAMOTO, A., FARRAR, W.T., & OVERBEEK, M. (1993). *Effects of grammatical category and stress pattern on on-line processing of bisyllabic words.* Unpublished 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., & 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., & 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. & JASSEM, W. (1965). Acoustic correlates of stress. *Language and Speech*, **8**, 159 – 181.

NAIGLES, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, **17**, 357 – 374.

NAKATANI, L. & 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., & 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. & JONGMAN, A. (1990). Phonological and form class relations in the lexicon. *Journal of Psycholinguistic Research*, **19**, 387 – 404.

SERENO, J.A. & JONGMAN, A. (1993). Acoustic correlates of form class. *Working Papers of the Cornell Phonetics Laboratory*, **8**, 119 – 140.

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

SOJA, N.N. (1992). Inferences about the meanings of nouns: The relationship between perception and syntax. *Cognitive Development*, **7**, 29 – 45.

SORENSEN, J., COOPER, W., & PACCIA, J. (1978). Speech timing of grammatical categories. *Cognition*, **6**, 135 – 153.

TREIMAN, R. (1989). The internal structure of the syllable. In G. Carlson & M. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 27 – 52). Dordrecht: Kluwer.

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

APPENDIX A

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	0.972	493	1.072	.515	1.160	.525	0.997	.453
2	1.155	.635	1.153	.478	1.080	.567	1.087	.498
3	1.290	.482	1.055	.470	1.188	.430	1.05	438
4	1.280	.577	1.317	.528	1.388	.560	1.25	.463
5	1.413	.538	1.272	.575	1.165	.570	1.093	.555

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	1.158	.960	1.107	1.085	1.068	.943	1.085	.947
2	1.190	.920	1.157	0.885	1.212	.925	1.147	.877
3	1.170	.943	1.120	0.965	1.210	.942	1.115	.945
4	1.230	.928	1.145	0.882	1.210	.897	1.212	.875
5	1.168	.933	1.150	0.925	1.130	.940	1.070	.975

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)

	Noun reading				*Verb reading*			
	Noun-dominant		*Verb-dominant*		*Noun-dominant*		*Verb-dominant*	
Speaker	*Fore*	*Back*	*Fore*	*Back*	*Fore*	*Back*	*Fore*	*Back*
1	0.847	0.855	0.812	0.858	0.825	0.870	0.842	0.845
2	1.210	0.980	1.058	0.967	1.170	0.928	1.160	0.905
3	1.340	1.075	1.265	1.097	1.408	1.097	1.355	1.040
4	1.043	1.225	1.008	1.093	1.092	1.170	1.028	1.035
5	1.783	0.925	1.553	1.048	1.538	0.975	1.202	1.220

APPENDIX B

For each speaker, mean values for duration (ms), amplitude (dB), and fundamental frequency (Hz) are presented for each syllable, for each Reading (noun reading, verb reading). Measurements are organized in terms of Dominance (noun-dominant and verb-dominant) and Stress (forestressed and backstressed). It should be noted that since a ratio of averaged values is not equivalent to an average of ratio values, these raw data means cannot be directly related to the mean ratio values listed in Appendix A.

		Duration (ms)		*Amplitude (dB)*		F_0 *(Hz)*	
	Speaker	*1st syllable*	*2nd syllable*	*1st syllable*	*2nd syllable*	*1st syllable*	*2nd syllable*
NOUN READING							
Noun-dominant forestressed	1	319	343	45.4	39.2	113	133
	2	294	296	37.6	31.7	121	100
	3	335	271	45.6	39.0	239	179
	4	385	314	44.3	36.0	185	178
	5	360	282	43.7	37.4	237	142
Noun-dominant backstressed	1	211	442	41.6	43.4	106	124
	2	237	401	34.9	38.0	110	113
	3	180	381	41.3	43.9	215	201
	4	226	404	37.0	39.8	204	169
	5	207	408	38.7	41.5	185	202

[continued overleaf]

		Duration (ms)		*Amplitude (dB)*		F_0 *(Hz)*	
	Speaker	*1st syllable*	*2nd syllable*	*1st syllable*	*2nd syllable*	*1st syllable*	*2nd syllable*
Verb-dominant forestressed	1	328	329	44.4	40.0	108	133
	2	304	292	37.4	32.4	118	115
	3	296	294	44.3	39.6	224	177
	4	384	311	41.4	36.2	184	183
	5	338	287	44.2	38.4	231	151
Verb-dominant backstressed	1	247	499	38.7	43.1	101	118
	2	214	460	33.1	37.4	107	110
	3	251	453	41.7	43.2	216	197
	4	237	489	37.2	42.3	198	181
	5	225	494	39.4	42.5	195	187
VERB READING							
Noun-dominant forestressed	1	336	316	45.7	42.8	107	130
	2	321	308	37.8	31.2	125	109
	3	332	293	47.0	39.1	248	176
	4	390	300	43.3	35.9	185	173
	5	325	278	44.5	39.3	240	156
Noun-dominant backstressed	1	212	424	43.3	45.9	97	113
	2	234	427	35.2	38.1	112	122
	3	173	402	41.6	44.1	211	194
	4	227	416	35.6	39.6	201	175
	5	201	370	39.1	41.6	188	195
Verb-dominant forestressed	1	314	329	47.2	43.7	107	127
	2	305	308	37.6	32.9	121	105
	3	282	278	45.5	40.8	232	173
	4	358	297	43.4	35.8	195	191
	5	325	302	42.6	39.8	189	179
Verb-dominance backstressed	1	210	493	43.6	46.0	91	109
	2	239	488	33.0	37.7	110	121
	3	190	456	41.4	43.7	207	199
	4	216	489	36.5	41.7	194	187
	5	228	438	40.8	41.8	201	169

Copyright of Language & Speech is the property of Kingston Press Ltd. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.