

DISAMBIGUATION OF DITROPIC SENTENCES: ACOUSTIC AND PHONETIC CUES

DIANA VAN LANCKER GERALD J. CANTER DALE TERBEEK
Northwestern University, Evanston, Illinois

In a previous study, we demonstrated that listeners were highly successful in identifying the intended meaning of spoken ditropic sentences (those which may carry either a literal or an idiomatic meaning) when speakers were instructed to convey the distinction. The present communication reports on acoustic and phonetic analyses carried out with the goal of identifying cues that distinguished the literal and idiomatic utterances. Certain prosodic differences were observed. Literal utterances were systematically longer than idioms. This was partly due to increased use of pauses, as well as to increased duration of major lexical items. Moreover, literal sentences were typically characterized by greater numbers of pitch contours (discernible rise-fall excursions of fundamental frequency) and open junctures than were idiomatic utterances. In addition to suprasegmental contrasts, articulatory distinctions—corresponding to lento-allegro phonological rules—were also observed.

These distinctions directly reflect the structural differences intrinsic to the two types of utterances. A literal sentence is formulated by the organization of constituent words and phrases. Idioms, on the other hand, are holistic units, largely nontransparent to syntactic structure or the usual meaning of the lexical members.

Ambiguity in language provides a challenge to investigators of speech and language behavior who have wondered how speakers and listeners deal with utterances that carry multiple meanings. Recent efforts in psycholinguistic research have used ambiguous sentences as a vehicle for investigating aspects of the processing of syntactic or semantic properties. (See Kess & Hoppe, 1979, for a current review of this work.) Most of these studies are concerned with sentences formed of constituents and differing in their internal structural configuration. For example, sentence 1 is structurally ambiguous as to whether its syntactic bracketing is represented by 2a or 2b:

1. The old men and women walked ahead.
- 2a. The [old [men and women]] walked ahead.
- 2b. The [[old men] and [women]] walked ahead.

The studies described in this paper focus on what we have termed *ditropically ambiguous* sentences. These are sentences that can have either an idiomatic or literal meaning, such as "She was to keep a stiff upper lip." (Was she to be brave, or was she to contract strongly the orbicularis oris muscle?)

The ambiguity of such sentences corresponds to the holistic/analytic dichotomy familiar to observers of language pathology. It is commonly observed in aphasia, for example, that patients can often utter an entire ready-made phrase with appropriate intonation contour, normal articulation, and normal temporal patterning but still be unable to produce normally the constituent parts of that phrase. One of our aphasic patients, for example, fluently produced the utterance "son of a bitch" but could not use the word *son* independently with its usual literal meaning. Such behavior illustrates how ditropic meanings are differentially organized. The idiomatic meaning

is carried by a unitary, unanalyzed phrase, the parts of which are not processed as constituents; the literal meaning, by contrast, is associated with a structurally formulated sentence.

This holistic/analytic distinction is further supported by psycholinguistic studies which have shown that the individual parts of idiomatic phrases are not stored or processed as discrete units, whereas the elements making up literal sentences are (Horowitz & Manelis, 1973; Osgood & Hoosain, 1974). Swinney and Cutler (1979) showed that subjects recognize idioms as natural English phrases faster than their literal counterparts. And Lieberman (1963) showed acoustic and perceptual differences between overlearned and novel sentences. We have previously reported experiments on the perception of ditropic sentences (Van Lancker & Canter, 1981). Five male speakers read aloud 15 pairs of sentences embedded in disambiguating paragraphs (see Appendix). When these recorded sentences were excised from context, listeners were not able to identify the intended meaning at better than chance level. Apparently speakers tend to allow semantic context to be the major cue for differentiating idiomatic from literal meanings. We found under certain conditions, however, that listeners are easily able to disambiguate ditropic sentence pairs even in the absence of a semantic context. In these experiments two male speakers produced 15 ditropic sentence pairs with instructions to convey as distinctly as possible the contrasting meanings. Each speaker then replicated his performance, yielding a total of 120 recorded sentences. Whether listening to these sentences in pairs or in unpaired random order, listeners identified sentence meaning correctly at a level far above chance (89.6% for the paired sentences and 85.5% for the randomized sentences).

In the present investigation, we used the same re-

corded sentences in an attempt to determine what acoustic cues had made possible such a high degree of discriminability. We analyzed the 120 recorded sentences by making measurements and observations of certain acoustic and phonetic parameters. The parameters were suggested by previous studies of the acoustic correlates of utterances which express contrasting meanings, either structural (Lehiste, 1973; Lehiste, Olive & Streeter, 1976; Scholes, 1971) or emotional (Fairbanks & Hoaglin, 1941; Williams & Stevens, 1972).

METHOD AND RESULTS

Sentence Duration

We first measured the total durations of the utterances from spectrographic records. The mean duration for the literal sentences (*L*) was 2.05 seconds ($SD = .62$), while the idiomatic sentences (*I*) had a mean of 1.54 seconds ($SD = .45$). The literal sentences overall were thus of longer durations (by one-third) than their idiomatic counterparts. This longer duration for literal sentences was characteristic of a large majority of the ditropic sentence pairs. In 52 of the 60 pairs, the literal member was longer than the idiomatic: 6 pairs were the same length; and only 2 pairs reversed the pattern. This distribution has a binomial probability of less than .001. Measures of speech phenomena contributing to sentence duration were needed to interpret this result.

Pauses, Open Junctures, and Pitch Contours

One of the factors contributing to duration differences is pausing. We found that there were typically more pauses in the literal than in the idiomatic members of the ditropic pairs. Compared to idiomatic sentences, literal sentences averaged about five times as many pauses, whether identified from spectrograms or from phonetic listening judgments. Pauses were measured from sound spectrograms by identifying gaps in the acoustic record not attributable to the articulation of stop consonant phonemes. In addition, we prepared two independent sets of phonetic transcriptions from listening to the tape-recorded sentences. Table 1 presents frequency counts of pauses taken from the spectrographic analysis and from phonetic transcriptions A (prepared by DT and DVL) and B (prepared by GJC) of the 60 literal and 60 idiomatic sentences.

The spectrographic results agreed well with both sets of phonetic judgments. In all three analyses, the occurrence of pauses was roughly five times greater in the literal sentences. One speaker's production of "She was to keep a stiff upper lip," illustrates clearly the role of pauses in differentiating between literal (*L*), with three pauses, and idiomatic (*I*) usage, with none.

3a. (*L*) She—was to keep a—stiff—upper lip.

3b. (*I*) She was to keep a stiff upper lip.

TABLE 1. Numbers of pauses and open junctures taken from the spectrographic analysis and from phonetic transcriptions (*A*, *B*, *C*) of 60 literal and 60 idiomatic sentences.

Source	Literal	Idiomatic
<i>Pauses</i>		
Spectrographic measurement	55	11
Phonetic judgment A	53	14
Phonetic judgment B	57	13
<i>Open Junctures</i>		
Phonetic judgment C	97	37

An additional phonetic analysis (*C*) focused on occurrences of open juncture, the phonetic phenomenon which signals linguistic (lexical or syntactic) boundaries (Lehiste, 1960). Open juncture is a perceptual category, indicated by a complex combination of acoustic and linguistic attributes. Pauses contribute to open juncture along with other acoustic variables, as illustrated by the data in Table 1, which show that the occurrence of distinct pauses falls far short of accounting for all the open junctures in our stimulus utterances. When we compared the frequency of open junctures in the two types of sentences, we observed the anticipated asymmetry. Open junctures, as shown in the table, occurred almost three times more often in the literal than in the idiomatic sentences.

Even without an associated pause, a drop in pitch can signal an open juncture. We therefore evaluated fundamental frequency patterns from spectrograms, looking for discrete rise-fall excursions, which we counted as a *pitch contour*. Figure 1 shows narrow-band spectrograms of a sentence pair, with the pitch contour traced along the tenth harmonic. Here the literal utterance contains two contours, whereas the idiomatic shows only one. But in other instances, these contours were not so distinct. We therefore made two counts of pitch contours, separated by four months. Both were based on judgments from unlabeled spectrograms. Table 2 shows some discrepancy in the counts from the first and second analysis. In both, however, the literal sentences included far more pitch contours than the idiomatic (from one-third to one-half again as many).

TABLE 2. Number of pitch contours taken from two analyses of spectrograms conducted four months apart.

Spectrographic analyses	Literal	Idiomatic
First analysis	115 ($\bar{X} = 1.92/\text{sentence}$)	78 ($\bar{X} = 1.30/\text{sentence}$)
Replicated analysis	138 ($\bar{X} = 2.30/\text{sentence}$)	105 ($\bar{X} = 1.75/\text{sentence}$)

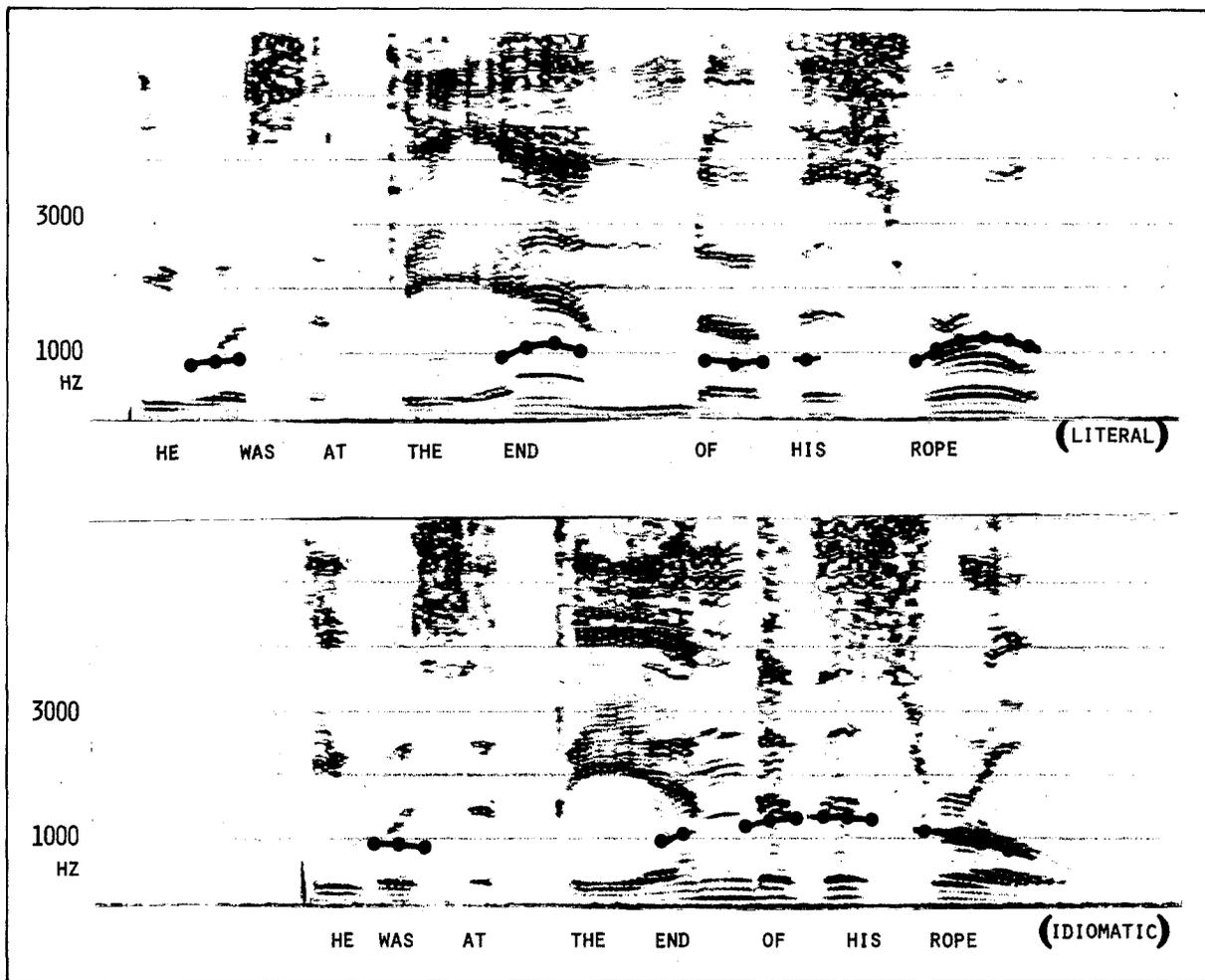


FIGURE 1. Narrow-band spectrograms of a ditropic sentence pair, with pitch contours traced along the tenth harmonic.

These findings are not at all surprising. In 1965, Bolinger demonstrated that “A accents”—such as that on

solid in *solid frame* (often associated with a pause)—signal syntactic “separateness,” whereas a contrasting kind of contour, a “B accent,”—such as that on

solid in *solid hour*—signals “connectedness.” We similarly observed that literal sentences are systematically marked by more such A-accent pitch contours and more pauses than idiomatic sentences, and these parallel our estimates of open junctures in these types of utterances. All of this is consistent with previous work showing the importance of prosodic parameters in signalling the details of the constituent structure of sentences (Lehiste & Wang, 1977; Nootboom, Brokx, & de Rooij, 1978; O’Malley, Kloker, & Dara-Abrams, 1973; Wingfield & Klein, 1971). There is little doubt that the frequent occurrence of these markers in literal sentences separates constituents intended to be understood as forming individual lexical items and phrases, whereas their absence contributes to the seamless shape of a holistic utterance.

Word Duration

Since longer durations of the literal sentences are partly accounted for by more pauses in these utterances, we wondered whether individual word durations might also contribute systematically to this pattern. For measurement we selected two words in each of the 15 sentences listed in the Appendix. Both target words were major lexical items at two different positions in the utterance: one located internally in the phrase and the other at the end of the phrase. For example, in “It broke the ice,” we measured *broke* and *ice*; in “She had him eating out of her hand,” we measured *eating* and *hand*.

We made spectrograms of the 120 total utterances (15 pairs spoken twice each by two speakers), then we compared the mean durations for medial and final words in both the literal and idiomatic utterances. Table 3 summarizes the data for this comparison for both speakers in their first and replicated readings.

It is evident that the mean word durations in the literal expressions are systematically greater than the corresponding words in the idiomatic expressions. The dif-

TABLE 3. Medial and final position mean word-duration contrasts (in milliseconds) of 120 literal and idiomatic utterances by two speakers (A and B) in two separate readings.

		Medial words				Final words			
		Literal	Idiomatic	<i>t</i> (<i>df</i> =28)	<i>p</i>	Literal	Idiomatic	<i>t</i> (<i>df</i> =28)	<i>p</i>
Speaker A									
First reading	\bar{X}	.544	.342	4.39	< .001	.510	.409	1.53	< .138
	<i>SD</i>	.109	.142			.209	.150		
Replicated reading	\bar{X}	.533	.322	7.90	< .001	.507	.439	2.03	< .062
	<i>SD</i>	.114	.098			.156	.156		
Speaker B									
First reading	\bar{X}	.474	.358	5.87	< .001	.542	.526	.711	< .489
	<i>SD</i>	.107	.090			.167	.149		
Replicated reading	\bar{X}	.422	.333	3.88	< .002	.579	.530	2.30	< .037
	<i>SD</i>	.116	.099			.183	.180		

ferences are greater for medial-position than for final-position words. All medial position differences are significant beyond the .01 level of confidence. However, in the final position, the difference is significant at the .05 level in only one of the four comparisons. This discrepancy can be explained by the phenomenon of prepausal lengthening, the tendency for syllables in phrase-final position to be lengthened (Gaitenby, 1965; Oller, 1973). This interpretation is consistent with studies showing that duration differences which might obtain in other positions are neutralized in the final position (Abercrombie, 1964). Probably some of the literal/idiomatic duration contrasts—with literal words tending to be longer—were masked by the tendency to lengthen *all* items in the prepausal position.

It was not unexpected that the individual content words of literal sentences would be typically longer than their idiomatic “twins” because the literal sentences tended overall to be longer. However, even in pairs whose total utterance lengths did not greatly differ, we found comparable word-duration differences. We selected all pairs whose literal/idiomatic (*L/I*) duration ratios were close to 1.0, in a range between the smallest ratio, .8 (literal slightly shorter), and 1.2 (literal slightly longer). The mean *L/I* ratio for these utterances was 1.09. We called these the “similar-length” pairs and compared their medial- and final-word durations with those in the remaining sentences (“different-length” pairs, with *L/I* ratios between 1.3 and 2.6, and a mean of 1.70). Information on the mean word durations from these two sets of sentences is presented in Figure 2. The expected literal/idiomatic differences in word duration are seen in the medial position, even for “similar-length” utterance pairs ($p < .01$). In the final position, however, no such differences were noted, likely because of the prepausal lengthening behavior mentioned earlier.

These results indicate that not only do individual word duration differences contribute to the sentence duration differences in literal versus idiomatic utterances, but that individual word durations (at least in a medial position)

may serve as cues for signalling literal/idiomatic meaning in utterances where overall sentence length is not a distinguishing feature.

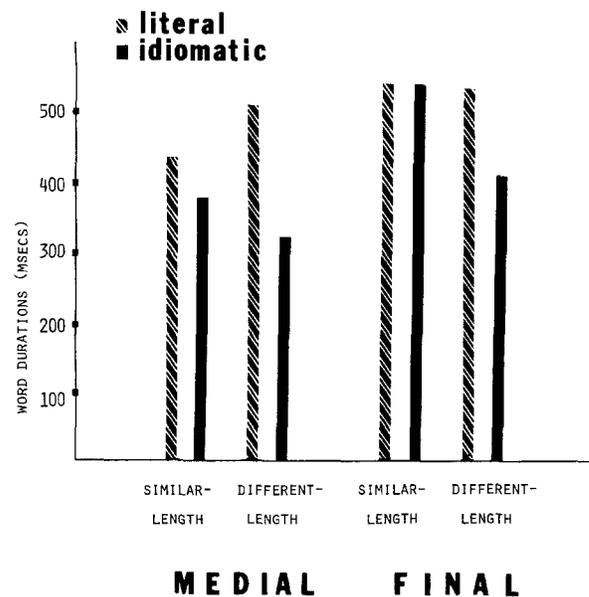


FIGURE 2. Medial and final mean word durations for “similar-length” and “different-length” sets of ditropic sentence pairs.

Articulation

Having considered prosodic differences that distinguish literal from idiomatic utterances, the next focus is on phonetic segment differences. Of course, these differences are intimately related to prosodic parameters such as duration and juncture. Our phonetic transcriptions revealed several articulatory behaviors contrasting in the ditropic sentence pairs, some of which are similar to differences between “allegro” and “lento” rules de-

scribed in phonological studies by Dressler (1972a, 1972b, 1972c) and Zwicky (1972), among others.

In general, the articulatory characteristics observed in the literal sentences were such that lexical and phrase boundaries were more strongly marked and individual lexical items were given increased salience. Stronger and oftentimes lengthened phonetic segments characterized these sentences. Segmental duration cues were undoubtedly used to signal both constituent structure and emphasis, as shown by Klatt (1975, 1976). The idiomatic sentences, by contrast, often showed more lax consonant and vowel production, tending to obscure lexical and phrase boundaries. In the sentence containing the phrase "an axe to grind," for example, both of our speakers initiated the word *axe* with a glottal stop in the literal expression but not in the idiomatic sentence. In the phrase "it's rotten to the core," both speakers produced perceptibly lengthened /r/'s in the key word *rotten* when the sentence was literal but not when it was idiomatic. These and like observations on segment lengthening support the proposal by Oller that word-initial and phrase-final lengthening effects may be cues for locating linguistic boundaries—words, phrases, and sentences. In the case of ditropic sentences, these effects do appear both for word-boundaries and for phrase boundaries. One of the speakers rather consistently deleted word-initial /h/ from unstressed words in the idiomatic sentences while retaining it in the literals. The second speaker sometimes produced the voiced cognate [ɦ]. (These behaviors were seen in such phrases as "eating out of *her* hand" and "out of *his* sails.") Both speakers on occasion used an aspirated stop [tʰ] in such words as *skating* and *eating* in their literal usage, but produced a tap [ɾ] in the same phonetic contexts when the phrases were idiomatic.

Vowel differences were generally in the direction of shorter, more neutral vowels for the idiomatic sentences, especially in unstressed words. The vowels in such words as *to*, *a*, and *the*, for example, were frequently reduced to [ə] in the idiomatic sentences. Diphthongs were sometimes reduced to pure vowels in the idioms, so that in the word "ice," *i* was produced as [aɪ] in the literal and as [a:] in the idiomatic utterance.

The two sentence pairs given below in phonetic transcription illustrate some of these articulatory contrasts in context:

4. He didn't know he was skating on thin ice.
 - 4a. (L) hi didnt nou | hi wAz s:ketʰɪŋ | ?ən θɪn | ?aɪs
 - 4b. (I) hi rɪdŋ no fi wAz | sketɪŋ ən θɪn a:s
5. She had him eating out of her hand.
 - 5a. (L) ʃi hæd him ?itʰɪŋ | ?aʊth əv hɜ hændʰ
 - 5b. (I) ʃi fiæd əm ɪrɪŋ aʊ əv ɜ fiændʰ

We do not present any frequency data on these observations because they were quite variable. Not only did our two speakers have some systematic stylistic differences, but one speaker often used different speech devices to mark his ditropic contrasts on different trials.

What seems quite evident is that speakers do use articulatory changes to help mark the appropriate meaning for ditropically ambiguous sentences and that various articulations may be substituted for one another to this end. Moreover, it is likely that similar trade-offs may occur between segmental and suprasegmental markers. Further research in which segmental and suprasegmental cues are systematically quantified will be required to specify the nature of these presumed interactions.

CONCLUSION

We undertook this series of acoustic measurements and phonetic analyses to determine what information speakers provide that allows listeners to determine whether the intended meaning of an ambiguous sentence is literal or idiomatic. Clearly numerous cues may help listeners to disambiguate these sentences. We have shown that prosodic differences (sentence duration, pausing, pitch contours, open junctures, and word duration) and segmental phonetic differences tend to mark utterances as either literal or idiomatic. And, likely, there are additional cues such as voice quality differences and other such nuances of sociolinguistic style (see Giles & Powesland, 1975) which we have not investigated. It is best perhaps to consider these acoustic and phonetic events as *potential* cues, normally available to speakers and recognizable by listeners. Considerable variability is observed both between and within speakers as to which particular cues are selected in a given utterance. Presumably, different combinations and degrees of cues have equivalent potency in signalling to listeners the intended meaning of a ditropic sentence.

All the investigated parameters seem to contribute to a common denominator of meaning disambiguation: for the literal sentences the acoustic cues serve to separate and highlight the constituent parts, while for the idiomatic sentences the acoustic cues tend to signal the melding of constituents, enveloping them in a seamless, unitary utterance. Because in a ditropic sentence the idiomatic meaning is the more likely, as our previous work has shown (Van Lancker & Canter, 1981), it may be that the speaker attempting to express the literal meaning selects those cues that mark the sentence as having its special, less likely meaning. These cues are, in general, those which disunite the usually melded components of the sentence.

Overall, acoustic and phonetic details seem to reflect directly the contrasting structural attributes of ditropically ambiguous sentences. Literal sentences are formed by integrating genuine constituents, each of which maintains its usual lexical content-meaning. Idioms, on the other hand, being relatively nontransparent by definition to syntactic structure or to the usual meanings of the lexical members, are to be apprehended as holistic language units. The physical shapes of the literal and idiomatic utterances investigated in this study thus are seen as direct reflections of their underlying linguistic form and function (cf. Bolinger, 1965, p. 57).

ACKNOWLEDGMENT

This investigation was supported in part by Training Grant 5 T32 NS 07100 from the National Institute of Health.

REFERENCES

- ABERCROMBIE, D. Syllable quantity and enclitics in English. In D. Abercrombie, D. B. Fry, P. A. D. MacCarthy, N. C. Scott, and J. L. M. Trim (Eds.), *In honour of Daniel Jones*. London: Longmans, Green & Co., 1964.
- BOLINGER, D. On certain functions of accents A and B. In I. Abe & T. Kanekiyo (Eds.), *Forms of English*. Cambridge: Harvard University Press, 1965.
- DRESSLER, W. *Methodisches zu allegro-regeln*, 2. Internationale Phonologietagung, Vienna, Austria, 1972. (a)
- DRESSLER, W. Phonologische Schnellsprecheregeln in der Wiener Umgangssprache. *Wiener Linguistische Gazette*, 1972, 1, 1-29. (b)
- DRESSLER, W. Allegroregeln rechtfertigen lentoregeln: Sekundäre phoneme des Bretonischen. Innsbruck: *Innsbrucker Beiträage zur Sprachwissenschaft*, 1972, Band 9. (c)
- FAIRBANKS, G., & HOAGLIN, L. W. An experimental study of the durational characteristics of the voice during the expression of emotion. *Speech Monographs*, 1941, 8, 85-90.
- GAITENBY, J. The elastic word. *Haskins Laboratories Status Report on Speech Research*, 1965, SR-2, 3.1-3.12.
- GILES, H., & POWESLAND, P. *Speech style and social evaluation*. New York: Academic Press, 1975.
- HOROWITZ, L. M., & MANELIS, L. Recognition and cued recall of idioms and phrases. *Journal of Experimental Psychology*, 1973, 100, 291-296.
- KESS, J. F., & HOPPE, R. A. Directions in ambiguity theory and research. In G. D. Prideaux (Ed.), *Perspectives in experimental linguistics*. Leiden, Netherlands: John Benjamin, 1979.
- KLATT, D. H. Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics*, 1975, 3, 129-140.
- KLATT, D. H. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 1976, 59, 1208-1221.
- LEHISTE, I. An acoustic-phonetic study of internal open juncture. *Phonetica*, 1960, 5 (Suppl.), 1-54.
- LEHISTE, I. Phonetic disambiguation of syntactic ambiguity. *Glossa*, 1973, 7, 2.
- LEHISTE, I., & WANG, W. S-Y. Perception of sentence boundaries with and without semantic information. In W. Dressler & O. Pfeiffer (Eds.), *Phonologica*, 1976. Innsbruck: Ernst Beckvar, 1977.
- LEHISTE, I., OLIVE, J. P., & STREETER, L. A. Role of duration in disambiguating syntactically ambiguous sentences. *Journal of the Acoustical Society of America*, 1976, 60, 1199-1202.
- LIEBERMAN, P. Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, 1963, 6, 172-187.
- NOOTEBOOM, S. G., BROKX, J. P. L., & DE ROOIJ, J. J. Contributions of prosody to speech perception. In W. J. M. Levelt & G. B. Flores d'Arcais (Eds.), *Studies in the perception of language*. Chichester, England: Wiley & Sons, 1978.
- OLLER, D. K. The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America*, 1973, 54, 1235-1247.
- O'MALLEY, M. H., KLOKER, D. R., & DARA-ABRAMS, B. Recovering parentheses from spoken algebraic expressions. *IEEE Transactions on Audio and Electroacoustics*, 1973, AU-21, 217-220.
- OSGOOD, C., & HOOSAIN, R. Salience of the word as a unit in the perception of language. *Perception and Psychophysics*, 1974, 15, 168-192.
- SCHOLES, R. J. On the spoken disambiguation of superficially ambiguous sentences. *Language and Speech*, 1971, 14, 1-11.
- SWINNEY, D. A., & CUTLER, A. The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior*, 1979, 18, 523-534.
- VAN LANCKER, D., & CANTER, G. J. Idiomatic versus literal interpretations of ditropically ambiguous sentences. *Journal of Speech and Hearing Research*, 1981, 24, 64-69.
- WILLIAMS, C. E., & STEVENS, K. N. Emotions and speech: Some acoustical correlates. *Journal of the Acoustical Society of America*, 1972, 52, 1238-1250.
- WINGFIELD, A., & KLEIN, J. F. Syntactic structure and acoustic pattern in speech perception. *Perception and Psychophysics*, 1971, 9, 23-25.
- ZWICKY, A. M. Note on a phonological hierarchy in English. In R. Stockwell & R. Macaulay (Eds.), *Linguistic change and generative theory*. Bloomington: Indiana University Press, 1972.

Received March 31, 1980

Accepted July 8, 1980

Direct reprint requests to Diana Van Lancker, Research Associate, Department of Linguistics, UCLA, Los Angeles, California 90024.

APPENDIX

Sentence Stimuli Used (idioms underscored)

1. You find it's rotten to the core.
2. She was to keep a stiff upper lip.
3. We saw that we had left no stone unturned.
4. It broke the ice.
5. It came straight from the horse's mouth.
6. He was at the end of his rope.
7. It took the wind out of his sails.
8. David spilled the beans.
9. I hit the sack.
10. That's a real snake in the grass.
11. The coast was clear.
12. He didn't know he was skating on thin ice.
13. I also had an axe to grind.
14. It was the usual procedure about once a week to wash their dirty linen in public.
15. She had him eating out of her hand.