

Resolution of lexical ambiguity by emotional tone of voice

LYNNE C. NYGAARD and ERIN R. LUNDERS
Emory University, Atlanta, Georgia

In the present study, the effects of emotional tone of voice on the perception of word meaning were investigated. In two experiments, listeners were presented with emotional homophones that had one affective meaning (happy or sad) and one neutral meaning. In both experiments, the listeners were asked to transcribe the emotional homophones presented in three different affective tones—happy, neutral, and sad. In the first experiment, trials were blocked by tone of voice, and in the second experiment, tone of voice varied from trial to trial. The results showed that the listeners provided more affective than neutral transcriptions when the tone of voice was congruent with the emotional meaning of the homophone. These findings suggest that emotional tone of voice affects the processing of lexically ambiguous words by biasing the selection of word meaning.

Spoken word recognition relies on an acoustic speech signal that carries information not only about the linguistic content of a talker's utterance, but also about speaker-specific characteristics, such as talker identity, social status, and psychological state. In particular, a speaker's intonation, prosody, vocal effort, and speaking rate can convey an enormous amount of information about his or her emotional state (see, e.g., Murray & Arnott, 1993; Sherer, Banse, Wallbott, & Goldbeck, 1991). Although both linguistic and non-linguistic properties of spoken language are crucial for successful linguistic interpretation, relatively little attention has been paid to the integration of these two types of information in research on spoken language processing. In the present study, the nature of this integration is investigated by examining the influence of emotional tone of voice on the resolution of lexical ambiguity.

It is well known that spoken language is filled with ambiguities at the lexical level. Words often have more than one meaning, and listeners must resolve this ambiguity to access the appropriate meaning of a given lexical item. Research in which this issue has been investigated has focused primarily on determining at what *point* during the course of spoken language processing semantic and sentential context serve to disambiguate lexical items (e.g., Moss & Marslen-Wilson, 1993; Paul, Kellas, Martin, & Clark, 1992; Simpson, 1994). The question of what *kind*

of context can serve to disambiguate or bias the selection of a particular word meaning has received relatively less attention. Although investigators have examined which aspects of sentential context serve to influence lexical selection (Kellas, Paul, Martin, & Simpson, 1991; Paul et al., 1992; Tabossi, 1988), contextual factors other than sentential context have not necessarily been examined systematically. It has been assumed, either implicitly or explicitly, that nonlinguistic aspects of spoken language are not necessarily relevant to lexical access and selection. On the one hand, researchers have investigated how the linguistic content of speech—the syllables, words, and sentences of speech—are processed and represented (e.g., Luce, Pisoni, & Goldinger, 1990; Marslen-Wilson & Warren, 1994; McClelland & Elman, 1986). On the other hand, researchers have investigated how listeners perceive the emotional state of a speaker from affective tone of voice (e.g., Frick, 1985; Murray & Arnott, 1993; Pittam & Scherer, 1993). Thus, these two areas of research have largely been considered separately (but see Friend, 1996; Wurm & Vakoč, 1996).

This assumption of independence stems in part from traditional explanations of linguistic processing and representation. Properties of the speech signal, such as emotional tone of voice, have been viewed as a source of noise that the perceiver must strip away or normalize to retrieve the abstract, canonical linguistic representations thought to underlie subsequent stages of linguistic processing (Halle, 1985; Joos, 1948; Kuhl, 1991, 1992; Shankweiler, Strange, & Verbrugge, 1977). According to this view, non-linguistic characteristics, such as tone of voice, should have little influence on linguistic processing. Contextual cues, such as emotional tone of voice, should not constrain lexical activation and selection.

Recently, however, research has suggested that variability in spoken language may not be stripped away during the early stages of speech perception but, rather, retained

Portions of this research were presented at the 134th Meeting of the Acoustical Society of America, San Diego, December 1997. The authors thank Audrey Katz, Anna Katz, Michael Koslowski, and Adam Warden for their assistance with data collection. In addition, we thank Arthur Samuel and Stephen Goldinger for helpful suggestions concerning revision. Correspondence concerning this article should be addressed to L. C. Nygaard, Department of Psychology, Emory University, Atlanta, GA 30322 (e-mail: lnygaard@emory.edu).

—Accepted by previous editorial team

and used (e.g., Pisoni, 1993). Several studies using both explicit and implicit memory tasks have demonstrated that variation in the speech signal due to talker's voice is retained in memory along with linguistic content (Bradlow, Nygaard, & Pisoni, 1999; Church & Schacter, 1994; Goldinger, 1996, 1998; Nygaard, Sommers, & Pisoni, 1995; Palmeri, Goldinger, & Pisoni, 1993; Schacter & Church, 1992). Schacter and Church, for example, have shown that listeners are sensitive to changes in talker's voice from study to test in implicit memory tasks, such as auditory stem completion. Listeners were more likely to complete an auditory stem with a word from the previously presented list if the stem was repeated in the same voice, as opposed to a different voice. Similarly, Palmeri et al. found that talker characteristics are retained in episodic memory as well. In a continuous recognition memory task, listeners were more likely to recognize a previously presented word if it was repeated in the same voice, as opposed to a different voice.

In addition to talker identity, other types of variation in the speech signal have been shown to influence word recognition as well. Studies have shown that intonation (Church & Schacter, 1994), speaking rate (Bradlow et al., 1999), and vocal effort (Nygaard, Burt, & Queen, 2000) all appear to be retained in lexical representations. Nygaard, Burt, and Queen found that recognition memory performance was superior when words were repeated with the same vocal effort than when repeated with a different vocal effort (e.g., shouted vs. whispered speech). Although these repetition effects were mediated by the judged typicality of a particular surface form type, it does appear that listeners are extremely sensitive to the perceptual characteristics of language.

The research on the retention of talker-specific characteristics in speech calls into question two basic assumptions of traditional *abstractionist* explanations of spoken language processing. The first is that the end product of perception is a series of abstract, canonical linguistic representations. Rather, this research suggests that lexical representations may be perceptually grounded. Goldinger (1998; see also Jusczyk, 1993) has proposed that lexical representations are exemplar based. Each instance of a previously heard word may be retained in memory and may contribute to subsequent word recognition. The second assumption is that nonlinguistic properties of speech are irrelevant with respect to the initial stages of language processing. Given the importance that nonlinguistic properties have in spoken communication and given that these properties are retained in long-term memory, these properties may influence linguistic processing. Nygaard, Sommers, and Pisoni (1994; Nygaard & Pisoni, 1998) found that listeners who learned the talker-specific characteristics of speech were better able to extract linguistic content. These results suggest not only that listeners learned specific nonlinguistic characteristics of speech, but also that this learning helped them recover the linguistic content.

Although the perceptual details of spoken words appear to be preserved and used during spoken language pro-

cessing, the question remains as to how and whether the information that these nonlinguistic attributes of speech provide is integrated into ongoing language processing and representation. Certainly, tone of voice provides the listener with an enormous amount of information about the emotional state and communicative intent of conversational partners. Listeners can identify particular emotions from a talker's voice (see, e.g., Frick, 1985; Pittam & Scherer, 1993), and difficulties with decoding nonverbal aspects of speech can result in social and interpersonal problems (see, e.g., Nowicki & Carton, 1997). Initially, however, the processing of emotional tone of voice and linguistic content do appear to be separate (Kitayama & Howard, 1994). Individuals with left-hemisphere damage often exhibit aphasia or difficulty with the comprehension and/or production of speech but show relatively less impairment for identifying and producing emotional prosody (Tucker & Frederick, 1989). Conversely, individuals with certain types of right-hemisphere damage have difficulty perceiving, categorizing, and producing emotional prosody but have relatively less impairment in resolving the linguistic structure of language (Bowers, Bauer, & Heilman, 1993). Other areas of the brain, such as the amygdala (Scott et al., 1997) and the hippocampus (Ghika-Schmid et al., 1997), have also been found to play a role in the processing of emotional tone of voice, particularly for negative emotions, such as fear and anger. In normal populations, studies have shown a right-hemisphere advantage for processing tone of voice and a left-hemisphere advantage for processing linguistic content (Buck, 1988; Joannette, Goulet, & Hannequin, 1990; Safer & Leventhal, 1977; Springer & Deutsch, 1998).

Although there appears to be an initial neural separation of the processing of emotional tone of voice and linguistic content, these two sources of information are integrated during the processing of spoken language. Our appreciation of linguistic devices, such as irony and sarcasm, rests on our ability to evaluate mismatches between linguistic content and prosodic contour. The assumption has been, however, that the integration of these two sources of information occurs after linguistic processing and the processing of emotional tone of voice have been completed. Consequently, current explanations of phonetic perception, lexical access, and syntactic parsing, for example, generally include little role for an effect of emotional tone of voice.

Recent research suggests, however, that linguistic prosody may influence both lexical and syntactic processing. Cutler (1994) has found that lexical prosody influences word segmentation, lexical activation, and lexical selection. Lexical stress appears to be included in phonological representations of spoken words and appears to help listeners identify word boundaries and to quickly access lexical items in memory. Similarly, Jusczyk, Cutler, and Redanz (1993) have found that infants are sensitive to the prosodic regularities of language and appear to use those regularities to initially segment the speech stream. Kjelgaard and Speer (1999) have found that sentential prosody appears to influence listeners' syntactic parsing behavior in such a way

that linguistic prosody can serve to resolve syntactic ambiguity. These findings suggest that multiple sources of information may be used by listeners to construct and process their initial representations of spoken language. Thus, we may need to expand our analysis of the types of information that serve to constrain the processing of spoken words. To that end, the present experiments were designed to evaluate whether a nonlinguistic property such as emotional tone of voice would influence the resolution of lexical ambiguity.

Emotional tone of voice was selected for study because it appears to be a particularly salient property of spoken language. A speaker's tone of voice colors our interpretation of a linguistic message, and listeners appear to interpret a talker's utterance in the context of the overall tone of voice. Recently, researchers have begun to suggest that emotion, including mood, facial expressions, and emotional tone of voice, may influence language processing (Halberstadt, Niedenthal, & Kushner, 1995; Kitayama, 1990, 1991, 1996; Niedenthal & Halberstadt, 1995; Niedenthal, Halberstadt, & Innes-Ker, 1999; Niedenthal & Setterlund, 1994; Niedenthal, Setterlund, & Jones, 1994). For instance, Kitayama and Howard (1994) found that emotional tone of voice appears to influence the interpretation of sentence-length utterances and, in turn, the emotional linguistic content of an utterance appears to influence listeners' judgments of emotional tone of voice.

With regard to the particular effects of emotional tone of voice on lexical access and selection, research results are mixed. Nygaard and Queen (2002; Nygaard, Queen, & Burt, 1998) found that emotional tone of voice appeared to influence the time course of lexical processing in an emotion-congruent manner. Nygaard and Queen presented listeners with emotion words produced with congruent, incongruent, or neutral emotional prosody. They found that listeners were faster to name or shadow words that were produced in a congruent tone of voice than those produced in an incongruent or neutral one. Listeners appeared to be integrating the speaker's tone of voice with linguistic content during spoken word recognition. Conversely, Wurm and Vakoch (1996) found little influence of emotional tone of voice on lexical decision times for words varying in emotional connotation. Although they found that the emotional connotation of a word influenced response times, the degree to which tone of voice was congruent or incongruent with respect to emotional meaning had little effect on responding.

Across areas of visual and auditory perception, the largest influences of top-down knowledge and context are observed when the input to the perceiver is ambiguous. For example, effects of such factors as lexical status or speaking rate on phonetic perception are generally found when incoming phonetic information is ambiguous or degraded (Ganong, 1980; Miller & Dexter, 1988; Nygaard, 1993; Samuel, 1996). The present experiments were designed to evaluate whether emotional tone of voice would influence the resolution of lexical ambiguity and the selection of word meaning. We reasoned that in instances in which lex-

ical meaning is ambiguous, as in the case of homophones, we would be most likely to find an influence of context in general and of emotional tone of voice in particular.

Our investigation consisted of two experiments. In both, listeners were asked to transcribe homophones that had one emotional meaning, either happy or sad, and one neutral meaning (e.g., bridal/bridle, die/dye). Because each homophone had a distinct spelling for the alternative meanings, we were able to infer which meaning was accessed from listeners' transcription performance. In addition, each homophone was produced in a happy, sad, or neutral tone of voice. We evaluated whether changing the tone of voice in which a homophone was produced would influence the selection of word meaning. Would the lexical selection process change depending on whether the emotional tone of voice was congruent, incongruent, or neutral with respect to the emotional meaning of each homophone? We predicted that emotional tone of voice would have an *emotion-congruent* effect on the selection of word meaning. That is, we expected that when the homophones with one happy and one neutral meaning were presented in a happy tone of voice, listeners would provide more happy transcriptions than when the homophone was presented in a neutral or sad tone of voice. Similarly, we expected that listeners would provide more sad transcriptions when the sad/neutral homophones were presented in a sad tone of voice.

EXPERIMENT 1

The first experiment was designed to determine whether emotional tone of voice would influence the selection of word meaning in lexically ambiguous homophones. Items were presented to listeners blocked by tone of voice, so that listeners heard the happy and sad homophones in only one tone of voice at a time. In this manner, we were able to compare transcription performance for the same homophones presented in each tone of voice. Transcription performance was evaluated because we reasoned that this task would allow us to assess the *content* and *products* of lexical resolution. Our aim was simply to determine whether emotional tone of voice, a nonlinguistic property of speech, could act as a strong constraint on a linguistic process, lexical resolution.

Method

Listeners. The subjects were 119 undergraduate students from Emory University. All were native speakers of American English and reported no known hearing or language disorder at the time of testing. They received partial course credit for their participation.

Stimulus materials. A list of 113 words was constructed on the basis of listeners' ratings of emotional meaning. Test items included emotional homophones that were selected to have one emotional meaning (happy or sad) and one neutral meaning. Filler items included neutral homophones with two neutral meanings and words with either a single emotional (happy or sad) or a neutral meaning.

Ratings of affective meaning were collected from 23 judges for all emotion-related and neutral homophones and words. Because all the homophones had a distinct spelling for each meaning, ratings of affective meaning were collected for one of the spellings from 11 judges

Table 1
Mean Ratings of Relatedness to Happy and Sad

Emotional Homophones	Happy	Sad
Happy (e.g., flower/flour)		
happy meaning	4.05	1.39
neutral meaning	1.58	1.19
Sad (e.g., die/dye)		
sad meaning	1.29	4.51
neutral meaning	1.44	1.32
Neutral (e.g., hair/hare)	1.36	1.18

and for the other spelling from 12 separate judges. All 23 judges rated the emotional and neutral filler words. The judges were asked to rate each word by using two separate 1- to 7-point Likert-type scales for its relatedness to happiness and for its relatedness to sadness (1, *not at all related*; 7, *highly related*). Word meanings were included according to the following criteria. Meanings selected as happy were rated at least one standard deviation above the mean for relatedness to happy and one standard deviation below the mean for relatedness to sadness. The opposite criteria were used for selecting sad meanings. Neutral meanings had ratings below the mean for both the happiness and the sadness scales. Happy homophones ($n = 12$) had one meaning rated as happy and one rated as neutral, using the above criteria. Sad homophones ($n = 10$) had one meaning rated as sad and one as neutral. Both meanings were rated as neutral for neutral homophones ($n = 9$). An additional 4 happy homophones were taken from Halberstadt et al. (1995), resulting in a total of 16 happy homophones. Filler words had a single meaning rated as happy ($n = 28$), sad ($n = 21$) or neutral ($n = 29$). Mean word frequency for the emotional meanings of happy and sad homophones did not differ significantly from the mean frequency of their neutral counterparts. In addition, mean word frequency for the happy, sad, and neutral filler words did not differ significantly from one another. Table 1 shows the mean ratings of relatedness to happy and sad for both meanings of the emotional homophones. The Appendix lists all the homophones and filler items used in this study.

All 113 words were recorded in each of three different tones of voice (happy, sad, and neutral) by one female and one male amateur actor. Words were read in random order in one tone of voice at a time. The actors were instructed to enunciate clearly, read at a comfortable rate, and not allow the meaning of the word to influence the tone of voice. Both spellings of the homophones were provided in order to discourage the speakers from pronouncing them differently depending on meaning. All the stimuli were digitally recorded in a sound-attenuated room with a Sony Digital Audio Tape-corder TCD-D7. Each stimulus item was redigitized on a PowerMac 7100/80 at

22.050-kHz sampling rate and was edited into separate files for presentation.

To confirm that the tone of voice manipulation was perceptually salient, tone of voice judgments were collected for all the stimuli. Thirty-six listeners were presented all 113 stimulus items in each of the three emotional tones and were asked to judge the tone of voice (happy, neutral, or sad) of each word. The words were presented individually in random order, and the listeners were instructed to ignore word meaning and to respond solely to the tone of voice. Any stimulus from either speaker that did not reach the criterion of 75% correct tone-of-voice judgments, averaged across listeners, was rerecorded, redigitized, and edited for presentation.

To confirm that the speakers did not pronounce the words differently depending on meaning, acoustic measurements were taken of each homophone in each tone of voice. Presumably, pronunciation of the homophones could differ, depending on the relationship between meaning and tone of voice. For example, speakers might have pronounced happy homophones in a "happier" tone of voice than they pronounced neutral or sad homophones. To ensure that this was not the case, duration, fundamental frequency (F_0), F_0 variation, and root-mean square (RMS) amplitude were measured for words in each condition and are shown in Table 2. Separate analyses of variance (ANOVA) were conducted for each acoustic measurement. Homophone type (happy, neutral, and sad) and tone of voice (happy, neutral, and sad) were factors. Each analysis revealed significant main effects only of tone of voice [duration, $F(2,200) = 42.39, p < .001$; F_0 , $F(2,200) = 13.94, p < .001$; F_0 standard deviation, $F(2,200) = 29.27, p < .001$; RMS amplitude, $F(2,200) = 7.93, p < .001$], indicating that each acoustic measure varied as a function of tone of voice. Main effects of meaning and interactions of tone of voice and meaning were not significant, indicating that the acoustic measures did not reliably differ as a function of meaning. Thus, the strength of the tone of voice manipulation appeared to be roughly comparable across homophone types.

Procedure. Homophone type (happy or sad) was manipulated as a within-subjects factor. Emotional tone of voice (happy, sad, or neutral) and the voice of talker (male or female) were manipulated as between-subjects factors. Each group of listeners was presented with the list of 113 words in only one tone of voice. Thirty-nine listeners heard the happy tone of voice, 40 listeners served in the sad tone-of-voice condition, and a separate group of 40 listeners served in the neutral tone-of-voice condition. Within each tone-of-voice condition, approximately half of the listeners were presented with the list produced by the male talker, and half were presented with the list by the female talker. Talker was manipulated to control for any effects of talker-specific idiosyncrasies.

The listeners were individually tested in a sound-attenuated room. The stimuli were presented binaurally over matched and calibrated

Table 2
Acoustic Measurements of Each Homophone Type
Produced in Each Tone of Voice Condition

Acoustic Measure	Tone of Voice	Homophone Type		
		Happy	Neutral	Sad
Duration (msec)	happy	745.12	714.13	706.68
	neutral	578.63	512.11	551.73
	sad	824.40	847.91	802.08
F_0 (Hz)	happy	203.15	208.12	203.08
	neutral	192.97	191.74	192.40
	sad	182.62	176.83	186.66
F_0 standard deviation (Hz)	happy	31.81	34.04	34.49
	neutral	27.92	25.70	22.27
	sad	19.69	21.39	19.31
Root-mean square amplitude (volts)	happy	.526	.631	.689
	neutral	.537	.552	.650
	sad	.452	.456	.436

Beyerdynamic DT100 headphones at approximately 75 dB (SPL). Stimulus presentation and data collection were controlled on line, using PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993) on a PowerMac 7100/80 computer. During the experiment, the words were presented in random order, and the listeners were instructed to type the word that they heard on the keyboard in front of them. After typing in their response, they were instructed to hit the return key to initiate the next trial. If the listeners did not respond within 4 sec, the next trial was automatically initiated. No information was given to the listeners about the presence of homophones. After the experiment was completed, the listeners were asked (1) whether they noticed the presence of homophones, (2) how many homophones they remembered hearing, (3) whether they noticed that some homophones had emotional meanings, (4) how they chose which spelling to report, and (5) whether they had ever started typing one spelling but then changed their minds. These questions were designed to determine whether the listeners were responding strategically on the basis of the demands of the experimental situation.

Because each meaning of the homophone had a unique spelling, it was assumed that the listeners had accessed the emotional meaning if they produced the emotional spelling. Likewise, if the listeners produced the neutral spelling, it was assumed that they had accessed the neutral meaning. Percentage of emotional transcriptions served as the dependent measure.

Results and Discussion

Figure 1 shows percentage of affective transcriptions, collapsed across speaker, plotted as a function of homophone type (happy/neutral and sad/neutral) and tone of voice. As the figure shows, the listeners were more likely to provide the happy transcription of the happy/neutral homophones when the words were produced in a happy tone of voice.

Likewise, the listeners provided more sad transcriptions of the sad/neutral homophones when they were produced in a sad tone of voice.

A 2×3 ANOVA, with homophone type (happy/neutral and sad/neutral) and tone of voice (happy, neutral, and sad) as factors, was conducted using percentage of emotional transcriptions, collapsed across speakers, as the dependent measure. The analysis revealed a significant main effect of homophone type [$F(1,116) = 35.95, p < .0001$], indicating that, overall, the listeners produced more affective transcriptions (relative to neutral transcriptions) for the sad homophones than for the happy homophones. As was predicted, the analysis also revealed a significant interaction between homophone type and tone of voice [$F(2,116) = 10.19, p < .0001$], indicating that the number of affective transcriptions varied as a function of emotional tone of voice. Post hoc Tukey's HSD comparisons revealed that, for happy/neutral homophones, the percentage of happy transcriptions was significantly greater in the happy tone of voice than in the neutral or sad tone of voice ($p < .05$). For sad/neutral homophones, the percentage of sad transcriptions was significantly greater in the sad tone of voice than in the neutral or happy tone of voice.

The listeners' responses after the experiment suggest that they were not responding strategically to the demands of the experimental task. In response to the question of whether they noticed the presence of homophones, all the listeners stated that they were aware of homophones. The

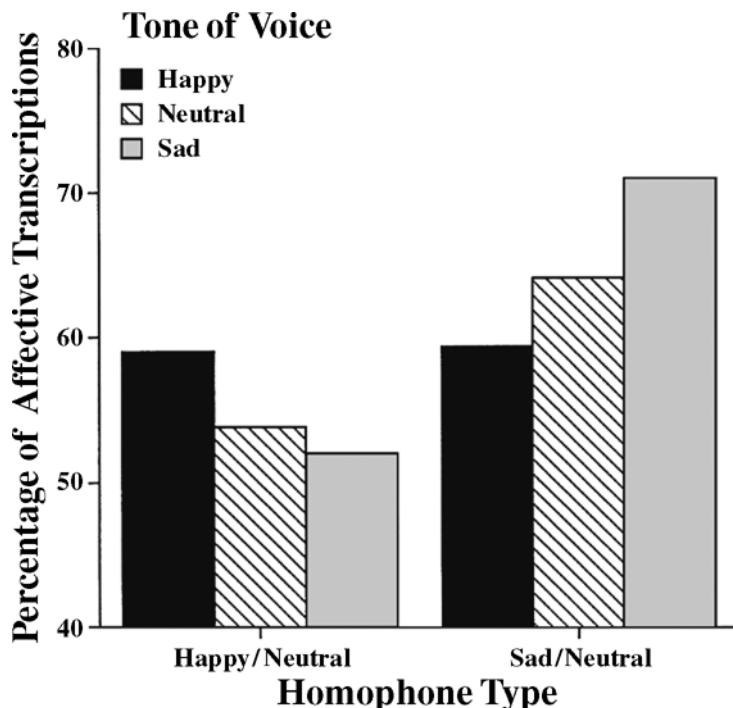


Figure 1. Percentages of affective transcriptions are plotted as a function of homophone type and tone of voice for the blocked tone-of-voice conditions of Experiment 1.

number of homophones they reported varied from 5 to 40. However, none of the listeners reported noticing that some of the homophones had emotional meanings. Ninety percent of the listeners indicated that they chose whichever spelling came into their mind first. However, 7% of the listeners indicated that the words preceding the homophone influenced which spelling they chose. Only 3 listeners mentioned that they started typing one spelling and changed it to the other. The listeners' responses indicated that although they appeared, in general, to be aware of homophones in the test list, they were not aware of meaning/tone-of-voice correspondences. Given that they did not report noticing homophones with one emotional and one neutral meaning, it is unlikely that they were responding to the explicit demand characteristics of the experiment.

These results indicate that the listeners were more likely to select the meaning of an ambiguous word that was congruent with the emotional tone of voice in which the word was produced. This finding suggests that the listeners not only were sensitive to the emotional tone of voice of the speaker, but also were integrating that tone of voice with the linguistic content of each word. The listeners appeared to use the emotional tone of voice to disambiguate the meaning of the emotional homophones. This finding is similar to Halberstadt et al.'s (1995) finding that listeners were more likely to choose an emotional meaning of an ambiguous homophone that was congruent with their induced mood. These findings are consistent with the view that listeners may be using emotional tone of voice as a type of context in which to guide their lexical selection.

Although emotional tone of voice appears to guide lexical selection, there are several ways in which it could have influenced the word recognition process. Because tone of voice was manipulated as a between-subjects factor, the listeners may not necessarily have been integrating tone of voice with their lexical processing of each word. Rather, the listeners may have adopted a general perceptual set consistent with the prevailing tone of voice and used it to guide lexical selection. This perceptual set could be either a general emotional expectation or even an induced mood resulting from a constant unvarying tone of voice. To evaluate these possibilities, the second experiment manipulated tone of voice as a within-subjects factor. From trial to trial, not only emotional meaning, but also emotional tone of voice varied. Thus, on any given trial, not only were the subjects unable to predict whether the meaning of each word was congruent, incongruent, or neutral with respect to tone of voice, but also they were unable to predict which tone of voice would be presented. The variation from trial to trial in tone of voice was assumed to prevent the development of a perceptual set or mood by the listeners. We predicted that tone of voice would influence the selection of word meaning in the absence of a perceptual set or mood.

EXPERIMENT 2

Method

Listeners. The subjects were 45 undergraduate students from Emory University. All were native speakers of American English and

reported no known hearing or language disorder at the time of testing. They received partial course credit for their participation.

Stimulus materials. Stimulus materials were the same 113 words as those used in Experiment 1. However, given the similar pattern of performance for the two speakers in the first experiment, only the female speaker's utterances were used. Three lists were constructed in which approximately one third of the happy, neutral, and sad homophones (and each type of filler word) was presented in a happy tone of voice, one third in a neutral tone of voice, and one third in a sad tone of voice. Lists were counterbalanced so that homophones were rotated through each tone of voice condition. Thus, across listeners, a given homophone was heard in each tone of voice.

Procedure. Homophone type (happy or sad) and emotional tone of voice were manipulated as within-subjects factors. Each group of 15 listeners was presented with one of the three counterbalanced lists of 113 words. Order of stimuli was randomized so that tone of voice and type of homophone varied from trial to trial. All other aspects of method and procedure were identical to those in the first experiment.

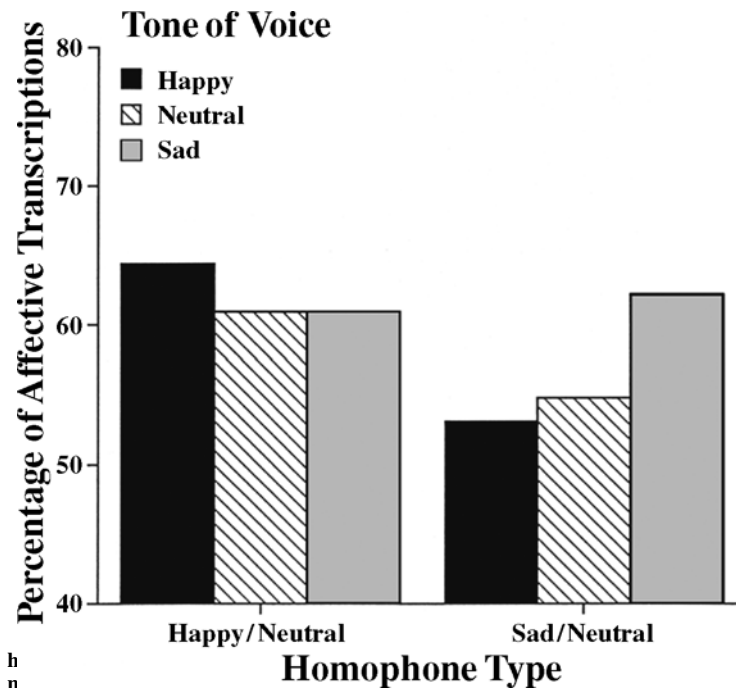
Results and Discussion

Figure 2 shows percentage of affective transcriptions plotted as a function of homophone type (happy/neutral and sad/neutral) and tone of voice. The listeners appeared more likely to produce the happy transcription of the happy/neutral homophones when the words were produced in a happy tone of voice and to provide more sad transcriptions of the sad/neutral homophones produced in a sad tone of voice.

A 2×3 ANOVA, with homophone type (happy/neutral and sad/neutral) and tone of voice (happy, neutral, and sad) as factors, was conducted using percentage of emotional transcriptions as the dependent measure. The analysis revealed a significant main effect of homophone type [$F(1,44) = 9.94, p < .005$], indicating that, overall, the listeners produced more affective transcriptions (relative to neutral transcriptions) for the happy homophones than for the sad homophones. This finding reflects the fact that the listeners produced more happy transcriptions, overall, for the female speaker and is consistent with the results of the first experiment. It is possible that the female speaker may have had, in general, a positive vocal affect, resulting in fewer sad and more happy transcriptions.

The analysis also revealed a marginally significant interaction between homophone type and tone of voice [$F(2,88) = 2.95, p < .06$], indicating that the number of affective transcriptions varied as a function of emotional tone of voice. Post hoc Tukey's HSD comparisons revealed that for happy/neutral homophones, the percentage of happy versus neutral transcriptions did not differ significantly across tone-of-voice conditions. For sad/neutral homophones, however, the percentage of sad versus neutral transcriptions was significantly greater in the sad tone-of-voice condition than in the neutral or happy tone-of-voice condition ($p < .01$).

Effect size estimates were calculated for the significant interactions in Experiments 1 and 2 between tone of voice and homophone type. In the first experiment, the interaction term accounted for approximately 17% of the variance (estimated $\omega^2 = .17$). In the second experiment, the interaction between these two factors accounted for ap-



proximately 4% of the variance (estimated $\omega^2 = .04$). According to Cohen's (1977) suggestions, the effect size for the interaction in the first experiment was fairly large, whereas the effect size in the second experiment was more moderate.

These findings suggest that although the effect of tone of voice was somewhat attenuated, relative to the first experiment, tone of voice did significantly influence the number of emotional transcriptions that the listeners produced. In particular, the listeners reliably chose the sad spelling of sad/neutral homophones when produced in a sad tone of voice. The difference between happy and sad emotions in this experiment has also been found in other tasks and domains. For example, Halberstadt et al. (1995) found, in a mood induction paradigm, that subjects provided more sad than neutral spellings for sad homophones when a sad mood was induced. However, subjects did not provide more happy than neutral spellings for happy homophones when a happy mood was induced. The reason for this asymmetry is unclear, but it is possible that our sad homophones in particular or sad emotional content in general were more salient or distinct from neutral content than were our happy homophones or happy emotional content in general. Despite the asymmetry, however, this finding does suggest that the listeners were not just adopting a perceptual set or developing a mood that was congruent with a particular tone of voice but, rather, that listeners appeared to use the emotional tone of voice of individual words to influence their lexical selection.

GENERAL DISCUSSION

In the present study, we investigated the role of emotional tone of voice in the resolution of lexical ambiguity. Two experiments were conducted in which listeners were asked to transcribe homophones that had both an affective and a neutral meaning. The homophones were presented with emotional prosody that was congruent, incongruent, or neutral with respect to the affective meaning. In the first experiment, tone of voice was blocked so that each group of listeners heard each type of homophone (happy and sad) in only one tone of voice (happy, sad, or neutral). In the second experiment, tone of voice was manipulated within subjects so that the listeners heard both types of homophones presented in all three affective tones. We found that, across experiments, the listeners provided more emotional transcriptions when the homophones were presented in a congruent tone of voice, as opposed to a neutral or incongruent tone of voice. These results suggest that the listeners selected meanings of ambiguous words that were congruent with the nonlinguistic properties of the speakers' utterances. Emotional tone appears to have been integrated during linguistic processing to constrain the listeners' selection of word meaning.

Voice expression was also shown to be related to word meaning in an experiment conducted by Kunihiro (1971). They found that English-speaking listeners were better able to guess the meaning of Japanese antonyms when the words were produced with an expressive voice than when

the words were read in a monotone or were presented visually. Taken together with our results, this finding suggests that tone of voice or emotional prosody not only may guide the selection of familiar lexical meaning, but also may be used to infer meaning from unfamiliar items. Listeners may use voice expressiveness or tone of voice in a manner similar to sentential context to constrain possible lexical candidates and to derive meaning from unfamiliar or noisy signals.

Although emotional tone of voice influenced lexical processing, it did not completely constrain the listeners' selections. In both experiments, the listeners still provided neutral transcriptions a certain percentage of the time. This finding suggests that emotional tone of voice may act as one of many possible constraints on lexical selection. Various constraints, such as the particular frequency of occurrence of a given word meaning (see Simpson, 1994, for a review of constraints on the resolution of lexical ambiguity), may each contribute differentially to the lexical activation and selection process to result in each listener's ultimate selection. For example, in the case of our emotional homophones, although mean word frequency was equated across emotional and neutral meanings, for any given homophone, one meaning may have been more frequent. In cases in which the neutral meaning was higher in frequency than the emotional meaning, it may have been more difficult to demonstrate an influence of emotional tone of voice. Just as is the case with sentential or linguistic semantic content, word frequency may have interacted with the effects of emotional tone of voice.

This finding that emotional tone of voice did not completely constrain transcription performance also suggests that tone of voice may serve as a somewhat weaker constraint than does sentential context during lexical processing. Given the off-line nature of our task, one would assume that any effects of a constraining context would have had plenty of time to be integrated with on-going lexical processing, at the activation, selection, or decision stages. That listeners still provided neutral transcriptions some of the time suggests that even in congruent tone-of-voice conditions, listeners still had the neutral meaning available at the time of transcription. In cross-modal priming tasks, research suggests that after as little as 600 msec, only the meaning that is consistent with a constraining sentential context is still activated (e.g., Paul et al., 1992). Although numerous differences exist between the processing demands of cross-modal priming and the current transcription task, emotional tone of voice does appear to influence lexical selection either in a different manner or to a different degree than does sentential context.

One way in which emotional tone of voice might influence lexical selection is that it may be preserved, along with linguistic content, in some kind of integrated lexical representation. For these homophones, the same phonological form points to two meanings, one emotional and one neutral. However, if tone of voice is preserved in memory and is correlated with a particular meaning in the listener's experience, listeners should choose the orthographic representation that reflects a consistent pairing of tone of voice

with a particular meaning. For example, for a happy homophone, a happy tone of voice may occur more often with the happy meaning than with the neutral meaning. Consequently, when the homophone is encountered in a happy tone of voice, the probability of accessing the happy meaning becomes greater.

This view has much in common with Goldinger's (1998) exemplar-based view. Goldinger (1998) suggested that spoken word recognition is based on an episodic lexicon. Aspects of a spoken word, such as a talker's voice and the context in which it occurs, were hypothesized to be included in the memory trace of every word. Each word is represented by a collection of instance-specific memory traces. On this view, perceptual and other characteristics of spoken words, such as voice and accompanying context, could all potentially contribute to the lexical access and retrieval process. On the basis of Hintzman's (1986, 1988) multiple-trace memory model, Goldinger's (1998) theory suggests that a to-be-recognized lexical item or probe activates a set of similar traces in memory. The intensity or strength of the activation, or echo, depends on the similarity of the probe to existing traces and to the number of traces in memory. If memory traces for words include such properties as voice identity or emotional tone, these properties, under the right circumstances, should influence lexical access and spoken word recognition.

In a series of experiments, Goldinger (1998) found that listeners' shadowing times, as well as the shadowing utterances themselves, reflected voice-specific properties of spoken words. For example, listeners' shadowing utterances mimicked the perceptual characteristics and, presumably, the acoustic characteristics of the spoken words that were shadowed. Furthermore, as the number of times an item was repeated in the same voice increased, the degree to which the shadowing utterance matched the original speaker's production increased as well. Within the conceptual framework of the model, when an item is repeated with the same voice, many traces will be created with those particular talker characteristics. When a probe with the same voice is subsequently presented, the similarity of the probe to a number of existing traces will be high, resulting in shadowing responses that preserve talker-specific characteristics.

With respect to homophones or words with multiple senses, Hintzman (1986) suggests that these types of items may be represented by nonoverlapping collections of memory traces. Each meaning of a homophone is hypothesized to be highly context dependent, so that only a context-appropriate sample of the memory traces for any given word form will be activated during processing. According to this view and Goldinger's (1998) theory and findings, it is possible that if spoken words are represented as instance-specific memory traces, homophones that had one emotional and one neutral meaning might have different nonverbal or tone-of-voice characteristics associated with each separate meaning. The emotional meaning of a homophone would have been previously encountered in specific contexts with a congruent emotional tone of voice. Encounters with the neutral meaning would be more variable with

respect to tone of voice. This view is consistent with our data and suggests that aspects of a spoken utterance, such as context and voice characteristics, are preserved in memory, along with linguistic content. Thus, the effects of emotional tone of voice on lexical processing do not result from a top-down or a postprocessing influence, but rather emerge as a direct consequence of the nature of lexical representation and memory retrieval.

An alternative possibility is that emotional tone of voice may act as a kind of semantic context during the selection of word meaning. According to this view, emotional tone of voice is perceived and processed separately from linguistic content but is integrated relatively early in the processing of spoken words. Although this task does not allow us to pinpoint precisely the stage at which emotional tone of voice is integrated, it does at least appear that tone of voice influences the lexical selection and decision stages of linguistic processing. Certainly, other research suggests that nonlinguistic properties of speech (e.g., Goldinger, 1998) and, in particular, emotional tone of voice (Nygaard & Queen, 2002) influence the time course of lexical processing. Consequently, tone of voice may serve to activate emotion-congruent representations in memory, thereby facilitating and influencing the processing of spoken words.

How emotional meaning is extracted and how emotion is represented in memory remains unclear. Some accounts suggest that emotion is represented categorically, so that particular primary emotions, such as happiness, sadness, or anger, constitute independent category structures (e.g., Bower, 1981, 1987; Niedenthal et al., 1999). Alternative accounts suggest that emotion is represented as values along continuous dimensions, such as valence (positive vs. negative) and/or arousal (high vs. low) (e.g., Feldman, 1995; Lang, Bradley, & Cuthbert, 1990). With respect to emotional displays, such as tone of voice, listeners may be either extracting primary categorical emotional content or, rather, evaluating tone of voice in terms of a dimensional analysis. Our stimuli were designed to be neutral with respect to these views. Our happy versus sad distinction maps onto a positive versus negative one. Nevertheless, in other studies in which the influence of emotion on aspects of cognition and perception was investigated, emotion appears to act in a category-specific manner (see Niedenthal et al., 1999).

In addition to the direct effects of tone of voice on lexical selection, there may also have been a more indirect influence of emotional tone of voice. Although the patterns of results across experiments were comparable, there did appear to be a somewhat larger effect of tone of voice in the blocked design than in the mixed design. Thus, in the first experiment, in addition to integrating tone-of-voice information with linguistic content, listeners may have been developing a perceptual set or expectation. The blocked design allowed the listeners both to fully perceive and to identify the particular tone of voice and to set up a set of expectations about the surface form of the next to-be-transcribed word. When no such set or expectation could be developed, as in the mixed presentation format of the

second experiment, the listeners were forced to reinterpret tone of voice on each trial, potentially leading to a somewhat attenuated effect of tone of voice. Interestingly, effects of a generally accepted lexical property, word frequency, also appear to be susceptible to differences in mixed versus blocked designs. For example, Whalen and Wenk (1993) reported that when homophones were read aloud, low-frequency spellings resulted in longer utterances, but only when blocked low-frequency and high-frequency lists were presented.

Although these experiments strongly suggest that emotional tone of voice constrains the lexical selection process, there are potential limitations to this study. One is that the listeners may have been sensitive to the demand characteristics of the task. They may have recognized that the words had alternative meanings that varied with respect to the speaker's tone and may have altered their responses accordingly. Aspects of the data argue against this conclusion. First, in Experiment 1, the listeners did not report noticing emotional homophones in the test list. Rather, the majority of listeners reported transcribing the first spelling that came to mind. Although not conclusive, these reports strongly suggest that the listeners did not recognize alternative meanings of the homophones and then respond accordingly. Second, if the listeners had noticed that our test list was studded with emotional homophones, we might have expected an overwhelming effect of tone of voice, limited only by the listeners' ability to perceive emotional prosody and identify the alternative meanings of the homophone. Given the number of nonhomophone distractors and given the graded nature of our effects, it seems likely that transcription performance reflected a real influence of tone of voice on the lexical selection process.

In summary, these findings call into question traditional accounts of spoken language processing. The linguistic system and, in particular, lexical access and spoken word recognition do not appear to be isolated from other types of auditory information. Nonlinguistic properties of speech are not necessarily irrelevant with respect to lexical processing. Rather, nonlinguistic aspects of spoken language, such as emotional tone of voice, appear to be highly relevant aspects of spoken language that are integrated into linguistic processing. Emotional tone of voice may act as one type of contextual constraint in lexical access and selection, either by virtue of its pairing with particular meanings or because listeners actually extract emotional meaning from tone of voice and then integrate it with linguistic content. This influence of emotional tone of voice suggests that current accounts of the nature of contextual constraints in the resolution of lexical ambiguity need to be reconsidered. Both linguistic and nonlinguistic properties, such as emotional tone of voice, appear to work together to constrain the eventual access and selection of word meaning.

REFERENCES

- BOWER, G. H. (1981). Mood and memory. *American Psychologist*, **36**, 129-148.
 BOWER, G. H. (1987). Commentary on mood and memory. *Behaviour Therapy & Research*, **25**, 443-455.

- BOWERS, D., BAUER, R. M., & HEILMAN, K. M. (1993). The nonverbal affect lexicon: Theoretical perspectives from neuropsychological studies of affect perception. *Neuropsychology*, *7*, 433-444.
- BRADLOW, A. R., NYGAARD, L. C., & PISONI, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception & Psychophysics*, *61*, 206-219.
- BUCK, R. (1988). *Human motivation and emotion* (2nd ed.). New York: Wiley.
- CHURCH, B. A., & SCHACTER, D. L. (1994). Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *20*, 521-533.
- COHEN, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press.
- COHEN, J., MACWHINNEY, B., FLATT, M., & PROVOST, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavioral Research Methods, Instruments, & Computers*, *25*, 257-271.
- CUTLER, A. (1994). The perception of rhythm in language. *Cognition*, *50*, 79-81.
- FELDMAN, L. A. (1995). Valence focus and arousal focus: Individual difference in the structure of affective experience. *Journal of Personality & Social Psychology*, *16*, 153-166.
- FRICK, R. W. (1985). Communicating emotion: The role of prosodic features. *Psychological Bulletin*, *97*, 412-429.
- FRIEND, M. J. (1996). From prosodic to paralinguistic function: Implications for affective development. *Dissertation Abstracts International*, *56*, 6417.
- GANONG, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception & Performance*, *6*, 110-125.
- GHIKA-SCHMID, F., GHIKA, J., VUILLEUMIER, P., ASSAL, G., VUADENS, P., SCHERER, K., MAIDER, P., USKE, A., & BOGOUSSLAWSKY, J. (1997). Bihippocampal damage with emotional dysfunction: Impaired auditory recognition of fear. *European Neurology*, *38*, 276-283.
- GOLDINGER, S. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *22*, 1166-1183.
- GOLDINGER, S. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*, 251-279.
- HALBERSTADT, J. B., NIEDENTHAL, P. M., & KUSHNER, J. (1995). Resolution of lexical ambiguity by emotional state. *Psychological Science*, *6*, 278-282.
- HALLE, M. (1985). Speculations about the representation of words in memory. In V. A. Fromkin (Ed.), *Phonetic linguistics* (pp. 101-114). New York: Academic Press.
- HINTZMAN, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*, 411-428.
- HINTZMAN, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*, 528-551.
- JOANETTE, Y., GOULET, P., & HANNEQUIN, D. (1990). *Right hemisphere and verbal communication*. New York: Springer-Verlag.
- JOOS, M. A. (1948). Acoustic phonetics. *Language*, *24* (Suppl. 2), 1-136.
- JUSZYK, P. W. (1993). From general to language specific capacities: The WRAPSA model of how speech perception develops. *Journal of Phonetics*, *21*, 3-28.
- JUSZYK, P. W., CUTLER, A., & REDANZ, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, *64*, 675-687.
- KELLAS, G., PAUL, S. T., MARTIN, M., & SIMPSON, G. B. (1991). Contextual feature activation and meaning access. In G. B. Simpson (Ed.), *Understanding word and sentence* (pp. 47-71). Amsterdam: North-Holland.
- KITAYAMA, S. (1990). Interaction between affect and cognition in word perception. *Journal of Personality & Social Psychology*, *58*, 209-217.
- KITAYAMA, S. (1991). Impairment of perception by positive and negative affect. *Cognition & Emotion*, *5*, 255-274.
- KITAYAMA, S. (1996). Remembrance of emotional speech: Improvement and impairment of incidental verbal memory by emotional voice. *Journal of Experimental Social Psychology*, *32*, 289-308.
- KITAYAMA, S., & HOWARD, S. (1994). Affective regulation of perception and comprehension: Amplification and semantic priming. In P. M. Niedenthal & S. Kitayama (Eds.), *The heart's eye: Emotional influences in perception and attention* (pp. 41-65). New York: Academic Press.
- KJELGAARD, M. M., & SPEER, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory & Language*, *40*, 153-194.
- KUHL, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, *50*, 93-107.
- KUHL, P. K. (1992). Psychoacoustics and speech perception: Internal standards, perceptual anchors, and prototypes. In L. A. Werner & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 293-332). Washington, DC: APA Press.
- KUNIHIRA, S. (1971). Effects of the expressive voice on phonetic symbolism. *Journal of Verbal Learning & Verbal Behavior*, *10*, 427-429.
- LANG, P. J., BRADLEY, M. M., & CUTHBERT, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review*, *97*, 377-395.
- LUCE, P. A., PISONI, D. B., & GOLDINGER, S. D. (1990). Similarity neighborhoods of spoken words. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing* (pp. 122-147). Cambridge, MA: MIT Press.
- MARSLÉN-WILSON, W., & WARREN, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review*, *101*, 653-675.
- MCCLELLAND, J. L., & ELMAN, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1-86.
- MILLER, J. L., & DEXTER, E. R. (1988). Effects of speaking rate and lexical status on phonetic perception. *Journal of Experimental Psychology: Human Perception & Performance*, *14*, 369-378.
- MOSS, H. E., & MARSLÉN-WILSON, W. D. (1993). Access to word meanings during spoken language comprehension: Effects of sentential semantic context. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*, 1254-1276.
- MURRAY, I. R., & ARNOTT, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America*, *93*, 1097-1108.
- NIEDENTHAL, P. M., & HALBERSTADT, J. B. (1995). The acquisition and structure of emotional response categories. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 33, pp. 23-64). San Diego: Academic Press.
- NIEDENTHAL, P. M., HALBERSTADT, J. B., & INNES-KER, A. H. (1999). Emotional response categorization. *Psychological Review*, *106*, 337-361.
- NIEDENTHAL, P. M., & SETTERLUND, M. B. (1994). Emotional congruence in perception. *Personality & Social Psychology Bulletin*, *20*, 401-410.
- NIEDENTHAL, P. M., SETTERLUND, M. B., & JONES, D. E. (1994). Emotional organization of perceptual memory. In P. M. Niedenthal & S. Kitayama (Eds.), *The heart's eye: Emotional influences in perception and attention* (pp. 87-113). New York: Academic Press.
- NOWICKI, S., JR., & CARTON, E. (1997). The relation of nonverbal processing ability of faces and voices and children's feelings of depression and competence. *Journal of Genetic Psychology*, *158*, 357-363.
- NYGAARD, L. C. (1993). Phonetic coherence in duplex perception: Effects of acoustic differences and lexical status. *Journal of Experimental Psychology: Human Perception & Performance*, *19*, 268-286.
- NYGAARD, L. C., BURT, S. A., & QUEEN, J. S. (2000). Surface form typicality and asymmetric transfer in episodic memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *26*, 1228-1244.
- NYGAARD, L. C., & PISONI, D. B. (1998). Talker-specific perceptual learning in speech perception. *Perception & Psychophysics*, *60*, 355-376.
- NYGAARD, L. C., & QUEEN, J. S. (2002). *Communicating emotion: Linking affective prosody and word meaning*. Manuscript submitted for publication.
- NYGAARD, L. C., QUEEN, J. S., & BURT, S. A. (1998). Effects of affective tone on spoken word recognition. *Proceedings of the 16th International Congress on Acoustics and 135th Meeting of the Acoustical Society of America*, *3*, 2061-2062.
- NYGAARD, L. C., SOMMERS, M. S., & PISONI, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, *5*, 42-46.
- NYGAARD, L. C., SOMMERS, M. S., & PISONI, D. B. (1995). Effects of stimulus variability on perception and representation of spoken words in memory. *Perception & Psychophysics*, *57*, 989-1001.

- PALMERI, T. J., GOLDINGER, S. D., & PISONI, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 309-328.
- PAUL, S. T., KELLAS, G., MARTIN, M., & CLARK, M. B. (1992). Influence of contextual features on the activation of ambiguous word meanings. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 703-717.
- PISONI, D. B. (1993). Long-term memory in speech perception: Some new findings on talker variability, speaking rate, and perceptual learning. *Speech Communication*, **13**, 109-125.
- PITTAM, J., & SCHERER, K. R. (1993). Vocal expression and communication of emotion. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 185-197). New York: Guilford.
- SAFER, M. A., & LEVENTHAL, H. (1977). Ear differences in evaluating emotional tones of voice and verbal content. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 75-82.
- SAMUEL, A. G. (1996). Does lexical information influence the perceptual restoration of phonemes? *Journal of Experimental Psychology: General*, **125**, 28-51.
- SCHACTER, D. L., & CHURCH, B. (1992). Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 915-930.
- SCOTT, S. K., YOUNG, A. W., CALDER, A. J., HELLAWELL, D. J., AGGLETON, J. P., & JOHNSON, M. (1997). Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature*, **385**, 254-257.
- SHANKWEILER, D. P., STRANGE, W., & VERBRUGGE, R. R. (1977). Speech and the problem of perceptual constancy. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Toward an ecological psychology* (pp. 315-345). Hillsdale, NJ: Erlbaum.
- SHERER, K. R., BANSE, R., WALLBOTT, H. G., & GOLDBECK, T. (1991). Vocal cues in emotion encoding and decoding. *Motivation & Emotion*, **15**, 123-148.
- SIMPSON, G. B. (1994). Context and the processing of ambiguous words. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 359-374). San Diego: Academic Press.
- SPRINGER, S. P., & DEUTSCH, G. (1998). *Left brain, right brain: Perspectives from cognitive neuroscience* (5th ed.). New York: Freeman.
- TABOSSI, P. (1988). Accessing lexical ambiguity in different types of sentential context. *Journal of Memory & Language*, **27**, 324-340.
- TUCKER, D. M., & FREDERICK, S. (1989). Emotion and brain lateralization. In S. Wagner & A. Manstead (Eds.), *Handbook of social psychophysiology* (pp. 27-70). New York: Wiley.
- WHALEN, D. H., & WENK, H. E. (1993, November). *Effect of the proper/common distinction on duration*. Paper presented at the 34th Annual Meeting of the Psychonomic Society, Washington, DC.
- WÜRM, L. H., & VAKOCH, D. A. (1996). Dimensions of speech perception: Semantic associations in the affective lexicon. *Cognition & Emotion*, **10**, 409-423.

APPENDIX

Emotional Homophones With Emotion Meaning Listed First and Neutral Homophones

Happy		Sad		Neutral	
ate/eight		banned/band		chews/choose	
bridal/bride		blue/blew		chord/cord	
caller/collar		bored/board		pair/pear	
dear/deer		die/dye		hair/hare	
flower/flour		groan/grown		hall/haul	
heal/heel		lone/loan		pause/paws	
hymn/him		pain/pane		heard/herd	
knows/nose		poor/pour/pore		meat/meet	
medal/metal		thrown/throne		toe/tow	
peace/piece		missed/mist			
petal/pedal					
presents/presence					
rose/rows					
sweet/suite					
tide/tied					
won/one					
Filler Words					
award	lucky	ache	shame	border	pottery
beauty	nice	blame	sorrow	chair	pump
cash	paradise	cancer	struggle	cycle	rice
charm	praise	crisis	upset	detect	rock
cheer	prize	depressed	worse	display	sheet
comedy	proud	disappoint	fence	sleeve	
dazzle	safe	distress	formula	soap	
enjoy	smile	empty	habit	spare	
flourish	sunny	failed	handle	stamp	
friend	triumph	gloomy	label	uniform	
funny	trust	grave	locate	switch	
glad	utopia	injury	marble	trace	
glory	losing	margin	track		
humor	morgue	noon			
kiss	poverty	panel			
laugh	regret	phone			