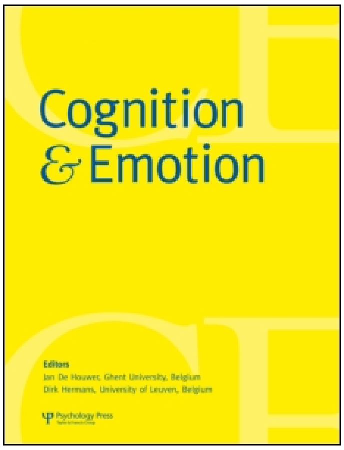

Cognition & Emotion
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713682755

Speech perception and vocal expression of emotion

Lee H. Wurm ; Douglas A. Vakoch ; Maureen R. Strasser ; Robert Calin-Jageman ;Shannon E. Ross

To cite this Article Wurm, Lee H. , Vakoch, Douglas A. , Strasser, Maureen R. , Calin-Jageman, Robert andRoss, Shannon E.(2001) 'Speech perception and vocal expression of emotion', Cognition & Emotion, 15: 6, 831 — 852
To link to this Article: DOI: 10.1080/02699930143000086
URL: http://dx.doi.org/10.1080/02699930143000086

COGNITION AND EMOTION, 2001, *15* (6), 831–852

Speech perception and vocal expression of emotion

Lee H. Wurm
Wayne State University, Detroit, USA

Douglas A. Vakoch
SETI Institute and University of California, Davis, USA

Maureen R. Strasser, Robert Calin-Jageman, and Shannon E. Ross
Wayne State University, Detroit, USA

Two experiments using identical stimuli were run to determine whether the vocal expression of emotion affects the speed with which listeners can identify emotion words. Sentences were spoken in an emotional tone of voice (Happy, Disgusted, or Petrified), or in a Neutral tone of voice. Participants made speeded lexical decisions about the word or pseudoword in sentence-final position. Critical stimuli were emotion words that were either semantically congruent or incongruent with the tone of voice of the sentence. Experiment 1, with randomised presentation of tone of voice, showed no effect of congruence or incongruence. Experiment 2, with blocked presentation of tone of voice, did show such effects: Reaction times for congruent trials were faster than those for baseline trials and incongruent trials. Results are discussed in terms of expectation (e.g., Kitayama, 1990, 1991, 1996) and emotional connotation, and implications for models of word recognition are considered.

There is no denying the salience of emotion or its importance in communicative contexts. Emotion colours our understanding of what is going on around us, our interpretation of the motives of people we interact with, and responses we deem to be appropriate. Organisms are provided with very powerful feedback about events in the environment through emotion (e.g., Buck, 1985; Hinde, 1974), and a person's emotional state can often be discerned even in the absence of any linguistic content (e.g., by observing someone's behaviour or body language). In the 1940s and 1950s, many researchers believed that the emotional content of a stimulus could be processed before the stimulus was consciously perceived. This was part of the so-called ''New Look'' in perception, which posited that per-

Correspondence should be addressed to Lee H. Wurm, Department of Psychology, Wayne State University, 71 W. Warren Avenue, Detroit, MI 48202, USA; e-mail: ae9075@wayne.edu

http://www.tandf.co.uk/journals/pp/02699931.html DOI:10.1080/02699930143000086

ception consists of finding partial clues, formulating a hypothesis about what the stimulus is, checking the data for verification, and then either accepting the hypothesis or reformulating it and trying again. The main idea was that organisms make inferences about category identity based on *partial* information from the environment, and an important corollary was that preconscious processing of emotion could influence subsequent perceptual processing. These ideas did not find their way into modern cognitive psychology for a number of reasons, and the New Look was largely swept away by the cognitive revolution (Kitayama and Niedenthal, 1994). However, some of the essential ideas of the New Look are making a reappearance in different areas of the literature.

Murphy and Zajonc's (1993) *affect primacy hypothesis* states that a nonconscious mental system processes affective information (e.g., whether something is seen as positive or negative). Evidence in support of this view comes from Bargh, Chaiken, Raymond, and Hymes (1996), who showed that words with similar valences primed each other in a naming task, while words with dissimilar valences did not. The word *flowers*, for example, primed the word *knowledge*, because both are judged to be positive (see also Bargh, 1990; Chartrand & Bargh, 1996; Fazio, 1990; Fazio, Sanbonmatsu, Powell, & Kardes, 1986). This early evaluative process is believed to be general and widespread (Bargh, Chaiken, Govender, & Pratto, 1992; Bargh et al., 1996). It seems to occur automatically and preconsciously for a wide range of objects, including objects that do not elicit a strong evaluative attitude. This notion fits quite naturally with the model of appraisal recently proposed by Robinson (1998; see also Scherer, 1984; Smith & Lazarus, 1990). In Robinson's model, the valence and urgency of a stimulus or situation are judged extremely rapidly and preattentively. The preattentive ''urgency detection'' module in the model determines whether a stimulus or event is personally relevant, whether it is consistent with or counter to a person's goals, and whether it can be effectively dealt with.

Although many studies have demonstrated selective attention to mood-relevant stimulus characteristics by psychiatric patients or mood-induced subjects (see, for example, Bower, 1981; 1987; MacLeod & Mathews, 1991; Mogg, Mathews, Eysenck, & May, 1991), the Bargh et al. (1992) study suggests that affect has a more pervasive role in perception. Their results fit with the conclusions we arrived at following a series of studies on spoken word recognition (Vakoch & Wurm, 1997; Wurm & Vakoch, 1996, 2000). These studies demonstrated that emotional connotation helps determine lexical decision times, not only for emotion words, but for words more generally. This conclusion held even after controlling for stimulus-based factors such as word frequency, word length, local phoneme-cluster characteristics, neighbourhood density, word onset characteristics, concreteness, imageability, familiarity, and animacy. Our results, like those of Bargh et al. (1992), suggest that affective information is processed very early in perception.

Converging evidence comes from several different experimental paradigms. Subjects in a study by Niedenthal (1990) were asked to identify cartoon characters previously seen in the experiment. Immediately prior to each trial, a human face was displayed as a prime for a very brief duration, and was not consciously detectable by the subjects. Despite the fact that subjects were not aware of the prime, they were faster to recognise the target cartoon character if the human face agreed in affect. This result suggests that nonverbal affective information can be perceived implicitly. Gernsbacher, Goldsmith, and Robinson (1992) found that emotion-appropriate probe words were named more quickly than emotion-inappropriate probe words, when subjects were reading stories that elicited emotional inferences. Finally, Williams, Mathews, and MacLeod (1996), using the ''emotional Stroop'' paradigm, found that anxious subjects were slower to name the colour of printed words if the words were threatening, and that this effect remained even when stimulus presentation was subliminal.

One more kind of evidence comes from research on the brain. Recent work has suggested that the categorisation of emotion takes place subcortically, and that parts of the brain responsible for perceptual processing are also implicated in the processing of emotion (e.g., Derryberry & Tucker, 1991, 1994; LeDoux, 1987, 1993). In addition, Hellige (1993) has shown that connotative information about objects can be transferred from one hemisphere to the other subcortically. These findings indicate that a link exists between perception and emotion, and complement existing behavioural data in suggesting that New Look ideas may be coming back.

There have been several studies having to do with emotion and language, many of which used mood induction (most commonly through music). Halberstadt, Niedenthal, and Kushner (1995) demonstrated that the resolution of lexical ambiguity can be influenced by emotion. They asked happy and sad mood-induced subjects to listen to homophones that had both happy and sad meanings, and to write down what they heard. Sad subjects wrote more sad spellings than did happy subjects, although happy subjects and sad subjects did not differ in their proportion of happy responses. Halberstadt et al. (1995) proposed a network model of language processing in which one form of top-down contextual feedback is the subject's emotional state.

Gerrig and Bower (1982) investigated whether subjects would identify words more quickly when the affective quality of the words matched the subject's emotional mood. Participants were induced to feel happy or angry, and the target words were pleasant, unpleasant, or neutral. Two lexical decision experiments failed to show any recognition advantage for mood-congruent words, and Gerrig and Bower concluded that mood is unlikely to affect lower-level pattern-identification processes in a highly overlearned, automatic task like word recognition. Niedenthal and Setterlund (1994; Niedenthal & Halberstadt, 1995) conducted a similar study but arrived at a different conclusion. In two lexical decision experiments, they found that RTs were faster for words related to the specific emotion induced (either happiness or sadness).

There have also been a number of studies that have used a tone-of-voice manipulation but in which no attempt was made to induce moods. These studies have attempted to exploit the fact that tone of voice is a valuable source of information that can be used by listeners to supplement purely linguistic aspects of messages (Murray & Arnott, 1993). Humour, cynicism, irony, and sarcasm work because the very same words can convey dramatically different messages, depending on the vocal expression of emotion with which they are spoken (Kitayama & Howard, 1994). One of the central questions of such studies has been whether the vocal expression of emotion is relevant perceptually or only at some later, post-perceptual processing stage.

The assumption (usually implicit) has been that speech and tone of voice are separate characteristics of messages that should be studied separately. Models of spoken word recognition have therefore not addressed the role of information conveyed by tone of voice, but have instead focused on abstract symbolic processing units that do not include any nonlinguistic features (Forster, 1979; Marslen-Wilson 1984; 1987; McClelland & Elman, 1986; Morton, 1969; Norris, 1994). A great deal of research with brain injured patients supports this separation, at least at some stages (e.g., Bowers, Bauer, & Heilman, 1993; Buck, 1988; Ghika-Schmid et al., 1997; Safer & Leventhal, 1977; Scott et al., 1997; Springer & Deutsch, 1985; Tucker & Frederick, 1989). In our first investigation of these issues (Wurm & Vakoch, 1996), we arrived at a conclusion similar to Gerrig and Bower's (1982). We investigated the relationship between emotional tone of voice of sentences and the emotional connotations of words, for subjects who did not undergo any kind of mood induction procedure. Although we did find that lexical decision times are affected by the emotional connotations of words, these effects were independent of the emotional tone of voice in which the stimulus sentences were presented. There was no advantage for simuli in which emotional connotation was congruent with the tone of voice, and there was no disadvantage for incongruent stimuli. We concluded that tone of voice is unlikely to affect something as automatic as speech perception.

However, some researchers believe that speech is an information-laden signal that far exceeds whatever formal linguistic message it contains. Exemplar-based models reflect this philosophy (Goldinger, 1998; Jusczyk, 1993), because the surface forms of utterances (including nonlinguistic aspects of the message) are considered to be important in such models. Evidence in support of this view comes, for example, from Lindfield, Wingfield, and Goodglass (1999). They examined identification of words that had been bandpass filtered so that the individual speech sounds making up each word could not be identified. The only information preserved by this method is the intonational contour of each word. The authors found that words could be reliably identified on the basis of the initial phoneme plus this intonational contour, which suggests that word prosody is represented in the mental lexicon. Also, stimulus variability frequently assumed to be extraneous (e.g., speaker gender, speaking rate, dialect) appears to be encoded during perception (Goldinger, 1996, 1998; Nygaard & Pisoni, 1998; Nygaard, Sommers, & Pisoni, 1994; Palmeri, Goldinger, & Pisoni, 1993; Pisoni, 1993). If something like the gender of the speaker is important enough to encode during perception, it is not difficult to imagine that tone of voice would be important enough, too.

Nygaard and Queen (1999) investigated this issue by measuring naming latencies for words presented auditorily in one of three tones of voice: happy, sad, or neutral. The semantic content of each of the words could similarly be described by those categories; a third of the words were happy words, a third sad words, and a third had neutral affective meaning. They found listeners were fastest to repeat words when the tone of voice in which the word was spoken matched the semantic category of the word. They concluded that emotional tone of voice and linguistic content are integrated very early in the recognition process with emotional tone of voice being seen as one kind of contextual information. Like other kinds of context effects, this information can be processed in parallel with the semantic content of words, and can serve to constrain lexical activation. Network models, too, could in principle incorporate tone of voice information with minimal modification. For example, feedback from a semantic level to lower perceptual levels could either speed or slow processing, depending on whether the tone of voice information was consistent or inconsistent with the meaning of a given word.

However, convincing proof that tone of voice can affect word recognition has been elusive. Furthermore, researchers even disagree about *how* emotional information would affect perception. Easterbrook (1959) argued that the primary action of affect would be to enhance cognitive processing, while Simon (1967) contended that affect would impair cognitive processing. Kitayama (1990, 1991, 1996) subsumes both possibilities under his notion of amplification (described more fully below), and he also noted that much of the disagreement in the literature may be due to subjects' expectations. In studies where subjects have a clear expectation of what is to be presented, affective information results in perceptual enhancement. In studies where such an expectation is not established, affective information results in perceptual impairment. Kitayama (1990) maintained that expectation can serve as a sort of preactivation of the memory code of the expected stimulus. However, with respect to something like speech perception, which Kitayama has not studied, it is a bit unclear what exactly would qualify as expectation.

Because one goal of the current study is to explore Kitayama's theoretical framework, it is important to be clear about what combinations of conditions predict what sort of perceptual performance. On one hand, perceptual impairment is predicted when subjects have a clear expectation and stimuli are impoverished or degraded. This is indeed what Kitayama (1991) found. He showed subjects positive, negative, or neutral words with extremely diminished contrast and brief exposure times. Subjects then indicated which of a pair of equally valenced words they thought was presented. Choice accuracy was lower for both positive and negative words, compared to neutral words, in Kitayama's (1991) view because in impoverished stimulus situations, ''. . . attentive processing is likely to be misdirected to an irrelevant perceptual code'' (p. 255). It is important for our present purposes to note that perceptual impairment is *not* predicted unless stimuli are degraded in some way. Experiment 1 of the current study was designed to provide this combination of non-impoverished stimuli and no clear expectation.

On the other hand, perceptual enhancement is predicted when subjects have a clear expectation and stimuli are not impoverished. This is because affective information can direct attention accurately to the relevant perceptual code. Experiment 2 of the current study is an attempt to observe this predicted perceptual enhancement, using nonimpoverished stimuli in an environment where an expectation is created.

The use of oral language is a more fundamental and widespread cognitive ability than reading and writing. Spoken language predates written language by at least 25,000 years, and probably much more than that (Pinker, 1994). Furthermore, all physically and mentally normal human beings learn to use spoken language, something that cannot be said of reading and writing. It therefore seems that if a link exists between language processing and emotion, it would be most evident with spoken language. From this perspective it is unfortunate that the majority of existing work in this area, including Kitayama's, has used visually presented materials (but see Kitayama, 1996). In the current study we use auditory stimulus presentation.

The present study has two related goals. One is to explore more fully the range of situations in which tone of voice affects lexical decisions. The second is to determine the extent to which a prevailing emotional context is consistent with Kitayama's idea of expectation. Experiments 1 and 2 of the present study used identical stimulus items, and differed only in how tone of voice was presented to subjects.

EXPERIMENT 1: WITHIN-SUBJECTS MANIPULATION OF TONE OF VOICE

Method

Participants. These were 92 undergraduate psychology students at Wayne State University. All were native speakers of English with normal hearing, and all received extra credit in a psychology course for their participation.

Materials. The words and pseudowords used in this study are listed in the Appendix. They are the same stimulus materials used by Wurm and Vakoch

(1996). Forty-eight words were selected from Morgan and Heise's (1988) study of pure emotion words, which are emotion words that are relatively free of cognitive and behavioural connotations (Clore, Ortony, & Foss, 1987; Ortony, Clore, & Foss, 1987). Four words were chosen from each of three semantic categories (Disgusted, Petrified, and Happy).

The remaining 36 stimuli were randomly chosen from the full list of emotion words. We did this so that our pool of items would not be concentrated near the endpoints of the dimensions of interest. Having items chosen only from the target categories could have made the purpose of the experiment more obvious to participants, and increased the likelihood of strategic responding.

Twelve of these latter 36 words were retained as filler items (see the Appendix). The other 24 were changed into pseudowords, which are required in order to make the lexical decision task meaningful (without them, the correct response is always ''Word''). We did this by changing the phoneme at the uniqueness point (UP) to a different phoneme from the same broad class (i.e., fricatives replaced fricatives, vowels replaced vowels, and so on). The UP is the point in the acoustic signal where the word in question diverges from all other words in English (see Marslen-Wilson, 1984). For example, if a person has only heard the partial word *por-*, then it is unclear what the word in question might be: it could be *pore*, *porous*, *porpoise*, *porch*, *porcelain*, or one of many others. One phoneme later (i.e., after hearing the [s] sound), the ambiguity is resolved: Only *porcelain* is consistent with the input. The [s] is the UP of *porcelain*.[1]

The emotion categories we chose are not as comparable to each other on word frequency as we might hope. Words from the Happy category had a higher mean frequency (43 occurrences) per million words) than those from the other categories (3 for Disgusted, 21 for Petrified; frequencies are from Francis and Kučera, 1982. However, the issue of word frequency is of limited importance in the present experiment because the identical words will be used in all conditions across all experiments. Our main interest was in using semantic categories that are maximally distinct in terms of denotation.

For the stimuli used by Wurm and Vakoch (1996), a graduate student with theatre experience read each stimulus item embedded in the carrier phrase: ''When that happened I felt ____''. The carrier phrase provided a natural and fluent prosodic context while keeping the potential effects of semantic context to a minimum. The carrier phrase was read four times for each stimulus item, once in each of the following tones of voice: Happy, Petrified, Disgusted, and Neutral. In total, there were 192 stimulus sentences, each of which was digitised

[1] Calculation of UPs is not as straightforward as it first appears. For example, the UP of *dog* depends on whether the listener also knows the word *dogma*. However, we are very familiar with the subject population that was used in the present study, and all of the words used in the present study were highly familiar words. Furthermore, our main comparisons of interest in this study involve RTs to the same words.

at a sampling rate of 10 kHz (low-pass filtered at 4.8 kHz) and stored in a disk file. These same disk files were used in the current study.

Manipulation checks. A total of 15 undergraduates at Wayne State University participated in a rating study, listening to the 192 stimulus sentences presented in a random order. The word or pseudoword that ended each sentence was digitally stripped from the carrier phrase, so that the tone-of-voice ratings would not be influenced by word meanings. Subjects pressed a button on a response box to indicate which emotion they thought was being portrayed by the tone of voice. Overall, the emotions conveyed by our sentences were classified as intended nearly two-thirds of the time (64%). Per cent agreement averaged 79.2% for Disgusted sentences, 54.4% for Petrified sentences, 53.4% for Happy sentences, and 68.2%[2] for Neutral sentences (chance agreement equals 25% for each sentence type). Our per cent agreement values for Petrified and Happy are quite comparable to those reported by Scherer, Banse, Wallbott, and Goldbeck (1991), who used a different rating methodology and five emotions instead of four. Scherer et al. (1991) found only 27.5% agreement for Disgust; it seems likely to us that this is due to the fact that their study also included Anger, while the current study did not.

We also made some objective measurements of the stimuli to add weight to the results of the rating study. We measured the rate of articulation (stimulus duration) and the intensity (root mean square amplitude) of each stimulus. Pairwise comparisons of the resulting values showed that Petrified stimuli had the highest intensity, followed by Disgusted, Happy, and Neutral. Rate of articulation was also highest for Petrified stimuli, followed by Happy, Disgusted, and Neutral ($p \leq .001$ for all adjacent comparisons, except that Disgusted stimuli had only marginally slower rate of articulation than Happy stimuli: $p = .07$).

These two manipulation checks suggest that the tones of voice differed in the ways intended (see Murray & Arnott, 1993), and that the emotions portrayed were recognisable by our subject population. Readers are also reminded that the identical stimuli will be used in both experiments reported here, so any short-comings the stimuli might have should apply equally to both experiments.

Procedure

Participants in the first lexical decision experiment were tested in groups of one to three in a sound-attenuating chamber. The 192 stimulus sentences were divided into four lists of 48 sentences. Each word was presented in exactly one tone of voice per list. For example, the word *scared* was heard in a Happy tone

[2] This includes one participant who classified *none* of the Neutral sentences as Neutral. Excluding this participant, per cent agreement for Neutral was 73.1%.

of voice in one list, in a Petrified tone of voice in another list, and so on. Participants were randomly assigned to hear one of the four stimulus lists. Within each list the order of presentation was randomised. Each group of one to three participants got a different random order.

Within each list, 25% of the stimuli were from each of the four tones of voice. Across the entire experiment, each word was heard in each tone of voice an equal number of times. For each participant, half of the items were words and half were pseudowords.

Digitised speech files were played for the participants over headphones at a comfortable listening level. Participants were directed to make a speeded lexical decision about the item in sentence-final position. Each participant used his/her dominant hand to make responses on a button board, pushing one button for words and another for pseudowords. Reaction times (RTs) were measured from the onset of the phoneme at the uniqueness point of each word. Measuring from this point is perhaps not ideal, as phonemes can vary dramatically in terms of how quickly after onset they can be identified (see Radeau, Mousty, & Bertelson, 1989). However, as with the case of word frequency mentioned above, our main comparisons of interest involve the identical stimuli (not just the same words, but the identical digitised speech files) across experiments.

Before the main experiment, participants heard a practice list of similar composition and performed lexical decisions. The practice list contained 12 sentences that were not used in the main experiment.

Experimental design and statistical analysis. We planned several kinds of analyses for this experiment. First, data will be categorised according to the 12 combinations of tone of voice and semantic category, and a repeated-measures ANOVA will be performed with semantic category and tone of voice as within-subjects factors.

In order to look specifically at the effects of congruence and incongruence, data will also be categorised according to whether there was a match or mismatch between the semantic content of the word and the tone of voice on a given trial. Trials in which these matched are called Congruent trials, while those on which there was a mismatch are called Incongruent trials. All trials in the Neutral tone of voice are considered Baseline trials. We will perform all three pairwise comparisons between these trial types using repeated-measures ANOVAs. Comparing Congruent to Baseline gives us a test of perceptual enhancement, while comparing Incongruent to Baseline gives us a test of perceptual impairment. This experiment uses nonimpoverished stimuli but does not establish a clear expectancy about emotional content. As we outlined above, in this kind of situation we do not expect to observe perceptual enhancement or perceptual impairment, according to Kitayama's framework.

Diagnostic tests showed that our RT data did not meet the sphericity assumption. All F-tests reported in this paper remain significant when the recommended Box adjustment is applied (Hays, 1994; Huynh & Feldt, 1976). We report the corrected degrees of freedom and Huynh-Feldt ε where relevant.

Results and discussion

Subjects who had error rates of .25 or higher, or mean RTs of 1000 ms or higher, were excluded from the analyses. Twelve subjects' data were excluded by these criteria. RTs were discarded for trials on which the lexical decision was made incorrectly (4.7% of the data). RTs were discarded from the dataset if they were more than two standard deviations above the mean for a particular subject.

Figure 1 shows mean RT as a function of semantic category and tone of voice. A repeated-measures ANOVA revealed a significant effect of semantic category:

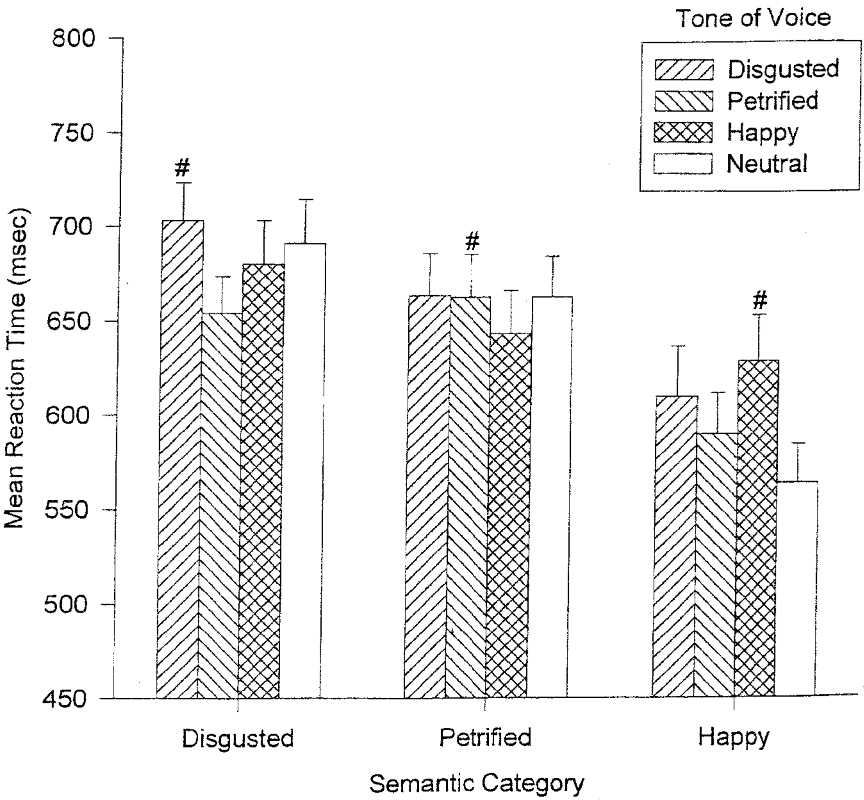

Figure 1. Mean reaction time (RT) as a function of tone of voice and semantic content, in milliseconds. "#" symbols show the conditions predicted to have relatively fast RTs. Error bars show +1 *SEM*. Data are from Experiment 1.

F(corrected $df = 1, 79) = 13.00$, Huynh–Feldt $\varepsilon = .5$, $p < .001$.[3] Participants responded more quickly to words belonging to the Happy semantic category than to those in the other two categories. This effect mirrors the one reported by Wurm and Vakoch (1996) and is probably due to the higher mean word frequency of those words. The main effect of tone of voice was not significant, nor was the interaction between tone of voice and semantic category. The ''#'' symbols in the figure indicate the conditions that should produce the shortest response latencies if there is any perceptual enhancement effect (i.e., the Congruent conditions). Inspection of the figure shows that these conditions did not produce fast RTs.

Mean RTs for Congruent, Incongruent, and Baseline trials were 658 ms ($SEM = 16.9$), 636 ms ($SEM = 13.9$), and 635 ms ($SEM = 16.9$), respectively. The pairwise comparisons showed that RTs did not differ as a function of trial type; none of the statistical tests approached significance. Mean error rates for Congruent, Incongruent, and Baseline trials were .033, .046, and .062, respectively. As with the RTs, none of the paired comparisons approached significance.

The results of this experiment indicate that hearing a word spoken in a tone of voice that is congruent with its meaning does not confer a lexical processing benefit. Hearing a word spoken in a tone of voice that is incongruent with its meaning does not lead to poorer lexical decision performance, either. Across two different types of RT analyses, there were no significant differences of any kind, except for the main effect of semantic category (which has nothing to do with the manipulations of interest). These results are consistent with the conclusion reached by Wurm and Vakoch (1996), and with the prediction of Kitayama's amplification model, discussed above. The next experiment used the same stimuli, but a different experimental design.

EXPERIMENT 2: BETWEEN-SUBJECTS MANIPULATION OF TONE OF VOICE

This experiment provided a test of the hypothesis that the vocal expression of emotion can affect perceptual processing when an expectation is created (Kitayama, 1990, 1991, 1996). The only change from Experiment 1 was in the presentation of the stimuli. Here, each subject heard all stimuli in the same tone of voice.

Method

Participants. These were 95 undergraduate psychology students at Wayne State University. All were native speakers of English with normal hearing, and all received extra credit in a psychology course for their participation.

[3] Condition means were substituted for critical trials on which the participant responded ''pseudoword'' (4.7% of the trials). This was necessary because each subject heard only one critical trial for each of the twelve specific combinations of tone of voice and semantic content.

Procedure

The procedure was identical to Experiment 1, except that the 192 stimulus sentences from Experiment 1 were divided into four lists different from those used above. In the present case, all sentences recorded in a given tone of voice were put on to the same list. As before, each word was heard exactly once by each participant. Participants were randomly assigned to hear stimuli in either the Disgusted, Petrified, Happy, or Neutral tone of voice. Within each list the order of presentation was randomised. Each group of one to three participants got a different random order.

Before the main experiment, participants heard a practice list of similar composition and performed lexical decisions. The practice list contained 12 sentences that were not used in the main experiment.

Experimental design and statistical analysis. We planned the same kinds of analyses as those we performed in Experiment 1. Data will be categorised according to the 12 combinations of tone of voice and semantic category, and an ANOVA will be performed. In this case, though, semantic category will be a within-subjects factor and tone of voice will be a between-subjects factor. As in Experiment 1, we will also perform all three pairwise comparisons between these trial types, analysing both RTs and error rates. The design of this experiment establishes a clear expectancy and uses nonimpoverished stimuli. According to Kitayama's theoretical framework, this should result in perceptual enhancement.

Results and discussion

The same exclusion criteria used in Experiment 1 were used here, resulting in the exclusion of 15 subjects' data. RTs for trials on which the lexical decision was made incorrectly were discarded. This occurred on 5.1% of the trials. As in Experiment 1, RTs were discarded if they were more than two standard deviations above the mean for a particular subject.

Figure 2 shows mean RT as a function of semantic category and tone of voice. We again conducted a repeated-measures ANOVA, this time with semantic category as a within-subjects factor and tone of voice as a between-subjects factor. The effect of semantic category was significant: F(corrected $df = 1.66,\ 126.15) = 37.23$, Huynh–Feldt $\varepsilon = .83$, $p < .001$, as in Experiment 1. The main effect of tone of voice was not significant, but in contrast to Experiment 1, the interaction between semantic content and tone of voice was significant: F(corrected $df = 4.98,\ 126.15) = 3.18$, Huynh–Feldt $\varepsilon = .83$, $p = .01$. The ''#'' symbols in the figure indicate the Congruent conditions. As the figure indicates, those conditions expected to show fast RTs did: For each semantic category, the fastest RTs were for words spoken in the corresponding tone of voice. This type of correspondence was completely absent from Figure 1.

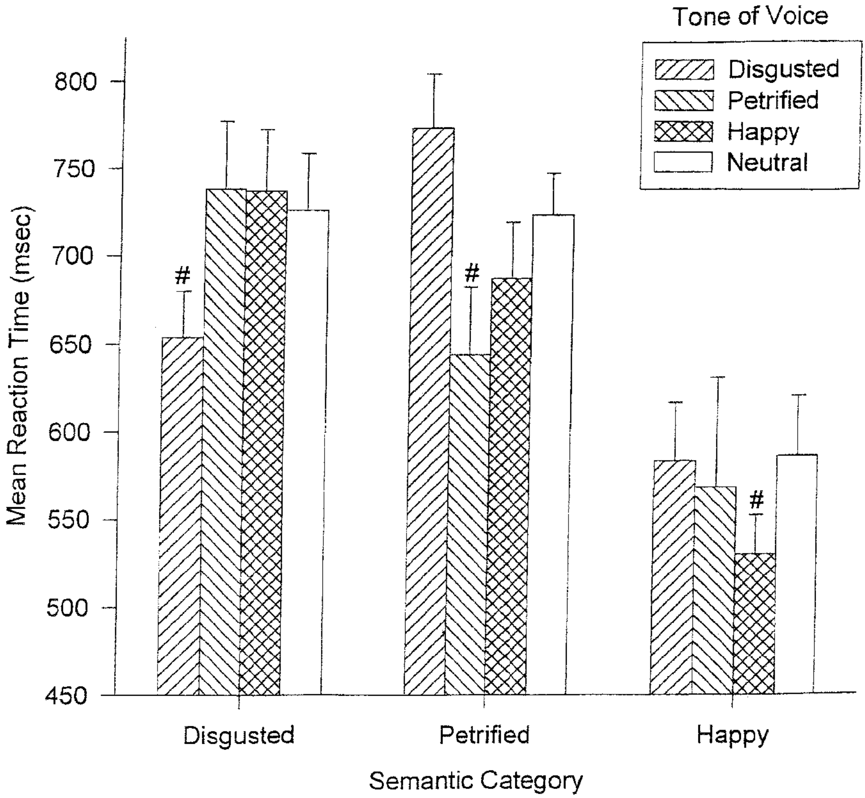

Figure 2. Mean reaction time (RT) as a function of tone of voice and semantic content, in milliseconds. ‘‘#’’ symbols show the conditions predicted to have relatively fast RTs. Error bars show +1 *SEM*. Data are from Experiment 2.

Mean RTs for Congruent, Incongruent, and Baseline trials were 609 ms ($SEM = 18.4$), 681 ms ($SEM = 18.0$), and 681 ms ($SEM = 26.0$), respectively. Pairwise comparisons of the RTs revealed a different pattern than that observed in Experiment 1. The RT advantage of Congruent trials over Baseline trials was significant: $F(1, 78) = 4.16$, $p < .05$, as was the advantage of Congruent trials over Incongruent trials: $F(1, 59) = 17.12$, $p < .001$. Incongruent RTs did not differ from Baseline. Mean error rates for Congruent, Incongruent, and Baseline trials were .046, .048, and .067, respectively. None of these differences approached significance.

These results stand in sharp contrast to those of Experiment 1, indicating that lexical decision performance is enhanced when the semantic content of the word matches the tone of voice of the carrier phrase. When a participant hears all stimuli spoken in the same tone of voice, a match between that tone of voice and the semantic content of the word facilitates lexical access.

We did some additional analyses to allow us to make this claim more definitively. Figure 3 shows the combined RTs for Experiments 1 and 2, excluding the Baseline trials.[4] We performed an ANOVA on the combined RTs, and found that the trial type $\times$ experiment interaction was highly significant: $F(1, 138) = 16.74$, $p < .001$. Follow-up comparisons showed that RTs on both trials types were significantly affected by the experimental manipulation: Congruent trials were significantly faster in Experiment 2 than in Experiment 1: $F(1, 138) = 3.83$, $p = .05$, while Incongruent trials were significantly slower in Experiment 2 than

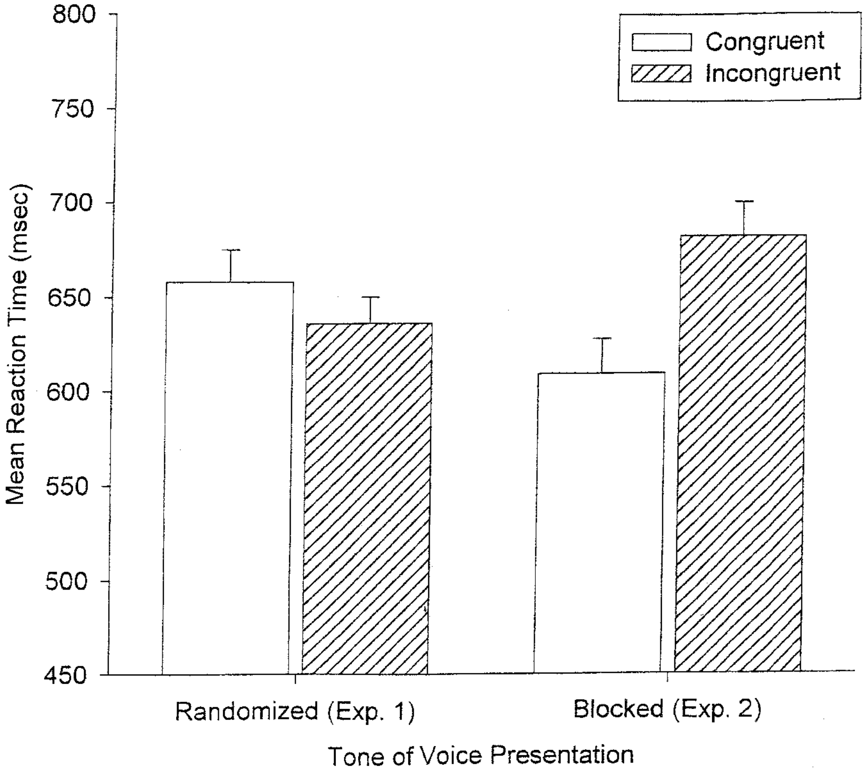

Figure 3. Mean reaction time (RT) as a function of trial type and experiment, in milliseconds. Error bars show +1 *SEM*.

[4] Inclusion of the Baseline RTs is not possible, because it would make trial type a between-subjects factor for some of the participants and a within-subjects factor for others.

in Experiment 1: $F(1, 138) = 4.00$, $p < .05$. These analyses confirm that the two experiments produced different results.

GENERAL DISCUSSION

We have previously attempted to demonstrate effects of congruence and incongruence between semantic content and tone of voice (Wurm & Vakoch, 1996), and we studied processing of semantic content in two additional studies (Vakoch & Wurm, 1997; Wurm & Vakoch, 2000). In those previous studies, we did not adequately take into account the way that this semantic processing is influenced by context (or expectation). With the present study we have examined the broader question of interaction between semantic content and the emotional context in which this content is presented.

This study demonstrates the importance of consistency of emotional context in determining whether a congruence effect is found, something that has not previously been demonstrated. A consistent emotional context was created in Experiment 2 through repetition of a single tone of voice, which may provide an example of what Kitayama (1990, 1991, 1996) has referred to as expectation. According to Kitayama's amplification model, when an expectation is created, perception is enhanced. In the present study, we observed a clear enhancing effect of congruence in the high expectation version of the experiment. It would be interesting to modify the current study, providing each subject with all four tones of voice but with particular words that come from must one of the three semantic categories. This would establish whether expectation can be created by semantic means, rather than prosodic means, and might have implications for the amplification model.

We see expectation as a low level (i.e., automatic, involuntary) aspect of perceptual processing (Bargh, 1990; Bargh et al., 1992; 1996; Chartrand & Bargh, 1996; Fazio, 1990; Fazio et al., 1986; Murphy & Zajonc, 1993). Therefore, even though it appears to develop over the course of multiple utterances, it should in principle be possible to establish it through the repeated presentation of words from the same semantic category. Of course, there would be the additional problem imposed by the definition of semantic categories themselves. When coming at these issues from the direction of prosody, it is relatively easy to proceed. However, it becomes difficult to see how such a large number of words belonging to one semantic category could be gathered (and how many might be enough?).

Also, in an experiment like the one being suggested, it would be difficult to separate any possible effects from a more general semantic priming effect, and to demonstrate convincingly that the effects are in fact perceptual rather than strategic. One could in fact ask if these alternate explanations might not apply to the current results, particularly because we used the lexical decision task which has been criticised for being contaminated by post-perceptual decision processes (e.g., Balota & Chumbley, 1984, 1985). Future research will undoubtedly shed more light on these issues, but we do not believe that a strategic explanation will be satisfactory. We base this conclusion in part on recent work we have done in the area of connotation and word recognition using not only the lexical decision task, but also the auditory naming task (Wurm & Vakoch, 2000; Wurm, Vakoch, Aycock, & Childers, 2001). This task is widely used in perception research, has withstood intense scrutiny as a perceptual measure, and is believed to be free of post-perceptual contamination.

The current study provides a new demonstration of a perceptual enhancement effect described in Kitayama's amplification model. On the Congruent trials in our Experiment 2, affect (via tone of voice) directs attention to a particular perceptual code (e.g., Disgust). We observed perceptual enhancement when the semantic content of the critical word activated the same perceptual code. This kind of effect is predicted by the amplification model, but was not found in Kitayama's (1991) study. In our view, the most likely explanation for this is that the ''nonimpoverished'' stimuli in his study were too difficult to perceive. Kitayama attempted to manipulate degree of impoverishment by altering contrast, but in both cases he used very brief exposure durations. It is possible that the durations were too brief for the stimuli to be considered completely non-impoverished. The current study equated perceptibility of the stimuli while varying only expectation.

It is likely that this also explains why we did not observe a processing cost for Incongruent trials. On these trials, the words do not trigger the same perceptual code as the tone of voice, and therefore one might have predicted perceptual impairment. However, as outlined in the Introduction, perceptual impairment of this kind is predicted only with impoverished stimuli (which is what Kitayama, 1991, observed). Our Incongruent trials were very different from this situation—both the semantic content and the tone of voice were quite readily identifiable, as indicated by the low overall error rates and by our manipulation checks. It is worth noting that Nygaard and Queen (1999), too, found perceptual enhancement but not perceptual impairment, in an experiment similar to our Experiment 2. Mean naming times for congruent *and* incongruent stimuli were faster than the baseline naming times (i.e., those spoken in a neutral tone of voice), although the differences were not always significant.

Experiment 2, with its repeated presentation of sentences in the same tone of voice, might be considered an acoustic variant of the Velten technique of mood induction (Velten, 1968). If this is the case, we may have unintentionally induced moods of Disgust, Fear, Happiness, or Neutrality in our participants, which leaves open the possibility of a different explanatory mechanism than the one we are proposing. However, we think it is unlikely that this happened for several reasons. First, participants in a speeded lexical decision experiment need to concentrate fairly hard on the task at hand, which has nothing to do with the tone of voice. The participants' task is focused on the portion of the stimuli carrying the semantic information, and this aspect of the stimuli was mixed and randomised in both experiments. Second, Gerrig and Bower (1982) conducted a study that is comparable to our Experiment 2, but which did include an intentional attempt to induce subjects' moods. They found no effect of congruence or incongruence in lexical decision performance. Finally, the testing session in the current study lasted five to six minutes, which is a much shorter period of time than is typically used in mood induction paradigms (e.g., Mukherji, Abramson, & Martin, 1982; Pignatiello, Camp, & Rasar, 1986; Velten, 1968).

The enhancement effect was larger in the current study when the emotion was negative. This is reminiscent of Kitayama's (1996) finding that negative but not positive emotions facilitated memory performance under a heavy cognitive load. Kitayama suggested two explanations for the asymmetry. The first was an acoustic explanation: Negative tones of voice have more abrupt changes in both pitch and intensity than positive voices do, and thus, they may induce more arousal. The second possibility offered by Kitayama was *automatic vigilance* (Pratto & John, 1991), which is the idea that negative stimuli can automatically summon more attention than positive stimuli (see also Easterbrook, 1959). This notion also fits with Robinson's (1998) appraisal model, which asserts that only negative emotions (specifically fear and anxiety) can be generated solely by unconscious processing. These possibilities are intriguing and worthy of additional investigation, but they would seem to predict faster overall RTs for the negative tones of voice (at least in the Congruent conditions), contrary to the results of the present study. Furthermore, as we just discussed, we do not believe that we induced moods in our subjects, or that perceiving an emotion is somehow equivalent with being in an emotional state.

Together with our previous papers (Vakoch & Wurm, 1997; Wurm & Vakoch, 1996, 2000), the current study suggests that emotion is relevant perceptually, when looked at in the proper light. The Congruent trials from Experiment 2 bear the most resemblance to what we might think of as natural speech: In general, tone of voice does not rapidly switch from sentence to sentence, but rather is at least broadly consistent over the course of multiple utterances. In addition, in natural speech the semantic content of an utterance typically matches the tone of voice of that utterance.[5] Examples of exceptions to this tendency include irony, sarcasm, and certain jokes, which work precisely because they are violations of the more general rule (Kitayama & Howard, 1994). The current study found enhancement effects only when conditions matched this general rule.

How might tone-of-voice information be integrated with semantic content during speech perception? Nygaard and Queen (1999), who found results very

[5] We do not know of any scientific studies that have demonstrated this, but we believe that perception would be much less efficient, and often confusing, if the two sources of information were frequently at odds with each other.

similar to those of Experiment 2 in the current study, view emotional tone of voice as one kind of contextual information that can constrain lexical activation. This operates much like other kinds of context effects (e.g., sentential context). In their view, emotional prosody and the semantic content of words are seen as two separate sources of information that are processed in parallel, and that interact. Our results suggest that this is not always the case—it appears only to hold when there is a consistent emotional expectation, like that created in Experiment 2.

As mentioned in the Introduction, models of spoken word recognition have not usually included a role for the information conveyed by tone of voice. Most rely on abstract symbolic processing units that do not include any nonlinguistic features (Forster, 1979; Marslen-Wilson, 1984, 1987; McClelland & Elman, 1986; Morton, 1969; Norris, 1994). As we pointed out, exemplar-based models may be more promising (Goldinger, 1998; Jusczyk, 1993). Such models do take into consideration aspects of the surface form of utterances, and presumably models such as these could easily be modified to accommodate the current results. An alternative approach is suggested by the architecture of network-style models. Most of the existing models of this type could be modified in a straightforward manner to incorporate tone of voice information as a kind of feedback from a semantic level to lower processing levels.

Manuscript received 3 April 2000
Revised manuscript received 5 February 2001

REFERENCES

Balota, D.A., & Chumbley, J.I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 340–357.

Balota, D.A., & Chumbley, J.I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language*, *24*, 89–106.

Bargh, J.A. (1990). Auto-motives: Preconscious determinants of social interaction. In E.T. Higgins & R.M. Sorrentino (Eds.), *Handbook of motivation and cognition* (Vol. 2, pp. 93–130). New York: Guilford Press.

Bargh, J.A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology*, *62*, 893–912.

Bargh, J.A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: Unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology*, *32*, 104–128.

Bower, G.H. 91981). Mood and memory. *American Psychologist*, *36*, 129–148.

Bower, G.H. 91987). Commentary on mood and memory. *Behaviour Research and Therapy*, *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.

Buck, R. (1985). Prime theory: An integrated view of motivation and emotion. *Psychological Review*, *92*, 389–413.

Buck, R. (1988). *Human motivation and emotion* (2nd ed.). New York: Wiley.

Chartrand, T.L., & Bargh, J.A. (1996). Automatic activation of impression formation and memorization goals: Nonconscious goal priming reproduces effects of explicit task instructions. *Journal of Personality and Social Psychology*, *71*, 464–478.

Clore, G.L., Ortony, A., & Foss, M.A. (1987). The psychological foundations of the affective lexicon. *Journal of Personality and Social Psychology*, *53*, 751–766.

Derryberry, D., & Tucker, D.M. (1991). The adaptive base of the neural hierarchy: Elementary motivational controls on network function. In R. Dienstbier (Ed.), *Nebraska symposium on motivation: Perspectives on motivation* (Vol. 38, pp. 289–342). Lincoln, NE: University of Nebraska Press.

Derryberry, D., & Tucker, D.M. (1994). Motivating the focus of attention. In P.M. Nidenthal and S. Kitayama (Eds.), *The heart's eye: Emotional influences in perception and attention* (pp. 167–196). New York: Academic Press.

Easterbrook, J.A. (1959). The effect of emotion on cue utilization and the organization for behavior. *Psychological Review*, *66*, 183–201.

Fazio, R.H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M.P. Zanna (Ed.), *Advances in experimental psychology* (Vol. 23, pp. 75–109). San Diego, CA: Academic Press.

Fazio, R.H., Sanbonmatsu, D.M., Powell, M.C., & Kardes, F.R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238.

Forster, K.I. (1979). Levels of processing and the structure of the language processor. In W.E. Cooper & E.C.T. Walker (Eds.), *Sentence processing: Psycholinguistics studies presented to Merrill Garrett* (pp. 27–85). Hillsdale, NJ: Erlbaum.

Francis, W.N., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston, MA: Houghton Mifflin.

Gernsbacher, M.A., Goldsmith, H.H., & Robertson, R.R. (1992) Do readers mentally represent characters' emotional states? *Cognition and Emotion*, *6*, 89–112.

Gerrig, R.J., & Bower, G.H. (1982). Emotional influences on word recognition. *Bulletin of the Psychonomic Society*, *19*, 197–200.

Ghika-Schmid, F., Ghika, J., Vuilleumier, P., Assal, G., Vaudens, P., Scherer, K., Maider, P., Uske, A., & Bogousslavsky, 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, and 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.

Hays, W.L. 91994). *Statistics* (5th ed.). New York: Harcourt Brace.

Hellige, J. (1993). *Hemispheric asymmetry: What's right and what's left*. Cambridge, MA: Harvard University Press.

Hinde, R.A. (1974). *Biological bases of human social behavior*. New York: McGraw-Hill.

Huynh, H., & Feldt, L.S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics*, *1*, 69–82.

Jusczyk, P.W. 91993). From general to language-specific capacities: The WRAPSA model of how speech perception develops. *Journal of Phonetics*, *21*, 3–28.

Kitayama, S. (1990). Interaction between affect and cognition in word perception. *Journal of Personality and Social Psychology*, *58*, 209–217.

Kitayama, S. (1991). Impairment of perception by positive and negative affect. *Cognition and 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 and S. Kitayama (Eds.), *The heart's eye: Emotional influences in perception and attention* (pp. 41–65). New York: Academic Press.

Kitayama, S., & Niedenthal, P.M. (1994). Introduction. In P.M. Niedenthal and S. Kitayama (Eds.), *The heart's eye: Emotional influences in perception and attention* (pp. 1–14). New York: Academic Press.

LeDoux, J.E. (1987). Emotion. In F. Plum (Ed.), *Handbook of physiology: Higher functions of the brain* (Vol. V, pp. 419–459). Bethesda, MD: American Physiological Society.

LeDoux, J.E. (1993). Emotional memory systems in the brain. *Behavioural Brain Research*, *58*, 69–79.

Lindfield, K.C., Wingfield, A., & Goodglass, H. (1999). The role of prosody in the mental lexicon. *Brain and Language*, *68*, 312–317.

MacLeod, C., & Mathews, A. (1991). Biased cognitive operations in anxiety: Accessibility of information or assignment of processing priorities? *Behaviour Research and Therapy*, *29*, 599–610.

Marslen-Wilson, W.D. (1984). Function and process in spoken word recognition. In H. Bouma & D.G. Bouwhuis (Eds.), *Attention and performance: Control of language processes* (Vol. X, pp. 125–150). Hove, UK: Lawrence Erlbaum Associates Ltd.

Marslen-Wilson, W.D. (1987). Functional parallelism in spoken word recognition. *Cognition*, *25*, 71–102.

McClelland, J.L., & Elman, J.L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86.

Mogg, K., Mathews, A., Eysenck, M.W., & May, J. (1991). Biased cognitive operations in anxiety: Artifact, processing priorities, or attentional search? *Behaviour Research and Therapy*, *29*, 459–467.

Morgan, R.L., & Heise, D. (1988). Structure of emotions. *Social Psychology Quarterly*, *51*, 19–31.

Morton, J. (1969). The interaction of information in word recognition. *Psychological Review*, *76*, 165–178.

Mukherji, B.R., Abramson, L.Y., & Martin, D.J. 91982). Induced depressive mood and attributional patterns. *Cognitive Therapy and Research*, *6*, 15–21.

Murphy, S.T., & Zajonc, R.B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology*, *64*, 723–739.

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. (1990). Implicit perception of affective information. *Journal of Experimental Social Psychology*, *26*, 505–527.

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: Advances in research and theory* (Vol. 33, pp. 23–64). New York: Academic Press.

Niedenthal, P.M., & Setterlund, M.B. (1994). Emotional congruence in perception. *Personality and Social Psychology Bulletin*, *20*, 401–410.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*, 189–234.

Nygaard, L.C., & Pisoni, D.B. (1998). Talker-specific perceptual learning in spoken word recognition. *Perception and Psychophysics*, *60*, 355–376.

Nygaard, L.C., & Queen, J.S. (1999). *Communicating emotion: Linking affective prosody and word meaning*. Manuscript submitted for publication.

Nygaard, L.C., Sommers, M.S., & Pisoni, D.B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, *5*, 42–46.

Ortony, A., Clore, G.L., & Foss, M.A. (1987). The referential structure of the affective lexicon. *Cognitive Science*, *11*, 341–364.

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, and Cognition*, *19*, 309–328.

Pignatiello, M.F., Camp, C.J., & Rasar, L.A. (1986). Musical mood induction: An alternative to the Velten technique. *Journal of Abnormal Psychology*, *95*, 295–297.

Pinker, S. (1994). *The language instinct: How the mind creates language*. New York: William Morrow.

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.

Pratto, F., & John, O.P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, *6*, 380–391.

Radeau, M., Mousty, P., & Bertelson, P. (1989). The effect of the uniqueness point in spoken-word recognition. *Psychological Research*, *51*, 123–128.

Robinson, M.D. (1998). Running from William James' bear: A review of preattentive mechanisms and their contributions to emotional experience. *Cognition and Emotion*, *12*, 667–696.

Safer, M.A., & Leventhal, H. (1977). Ear differences in evaluating emotional tones of voice and verbal content. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 75–82.

Scherer, K.R. (1984). Emotion as a multicomponent process: A model and some cross-cultural data. In P. Shaver (Ed.), *Review of personality and social psychology* (Vol. 5, pp. 37–63). Beverly Hills, CA: Sage.

Scherer, K.R., Banse, R., Wallbott, H.G., & Goldbeck, T. (1991). Vocal cues in emotion encoding and decoding. *Motivation and Emotion*, *15*, 123–148.

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.

Simon, H.A. (1967). Motivational and emotional controls of cognition. *Psychological Review*, *74*, 29–39.

Smith, C.A., & Lazarus, R.S. (1990). Emotion and adaptation. In L.A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 609–637). New York: Guilford Press.

Springer, S.P., & Deutsch, G. (1985). *Left brain, right brain*. New York: W.H. Freeman.

Tucker, D.M., & Frederick, S. (1989). Emotion and brain lateralization. In H. Wagner & A. Manstead (Eds.), *Handbook of social psychophysiology* (pp. 27–70). New York: Wiley.

Vakoch, D.A., & Wurm, L.H. (1997). Emotional connotation in speech perception: Semantic associations in the general lexicon. *Cognition and Emotion*, *11*, 337–349.

Velten, Jr., E. (1968). A laboratory task for induction of mood states. *Behaviour Research and Therapy*, *6*, 473–482.

Williams, J.M.G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, *120*, 3–24.

Wurm, L.H., & Vakoch, D.A. (1996). Dimensions of speech perception: Semantic associations in the affective lexicon. *Cognition and Emotion*, *10*, 409–423.

Wurm, L.H., & Vakoch, D.A. (2000). The adaptive value of lexical connotation in speech perception. *Cognition and Emotion*, *14*, 177–191.

Wurm, L.H., Vakoch, D.A., Aycock, J., & Childers, R.R. (2001). *Semantic effects in lexical access: Evidence from single-word naming*. Manuscript submitted for publication.

APPENDIX

Words
Disgusted category: annoyed, disgusted, displeased, irked
Petrified category: afraid, petrified, scared, terrified
Happy category: contented, glad, happy, pleased
Filler words: agitated, apprehensive, deflated, delighted, empty, fearful, outraged, overjoyed, passionate, resentful, shaken, unhappy

Pseudowords
charned, cheerlace, deprussed, downhorted, eloted, emballassed, exsouted, frighkened, horrisied, hurk, irete, jealoos, joywess, lonefome, lonery, lovezhick, mab, melantoly, micherable, overyelmed, prud, regletful, reliejed, sickemed

Note: Presentation of all stimuli was auditory. Items were heard at the end of the carrier phrase ‘‘When that happened I felt ____’’.