Published in final edited form as:

Cogn Emot. 2009 January 1; 24(7): 1133–1152. doi:10.1080/02699930903247492.

There's more to emotion than meets the eye: A processing bias for neutral content in the domain of emotional prosody

Lauren Cornew,

University of California, San Diego, CA, USA

Leslie Carver, and

University of California, San Diego, CA, USA

Tracy Love

University of California, and San Diego State University, San Diego, CA, USA

Abstract

Research on emotion processing in the visual modality suggests a processing advantage for emotionally salient stimuli, even at early sensory stages; however, results concerning the auditory correlates are inconsistent. We present two experiments that employed a gating paradigm to investigate emotional prosody. In Experiment 1, participants heard successively building segments of Jabberwocky “sentences” spoken with happy, angry, or neutral intonation. After each segment, participants indicated the emotion conveyed and rated their confidence in their decision. Participants in Experiment 2 also heard Jabberwocky “sentences” in successive increments, with half discriminating happy from neutral prosody, and half discriminating angry from neutral prosody. Participants in both experiments identified neutral prosody more rapidly and accurately than happy or angry prosody. Confidence ratings were greater for neutral sentences, and error patterns also indicated a bias for recognising neutral prosody. Taken together, results suggest that enhanced processing of emotional content may be constrained by stimulus modality.

Keywords

Emotion; Emotional prosody; Auditory emotion processing; Emotion bias

Perceiving emotional expressions is an essential component of social interaction and confers survival value by enabling us to discriminate friend from foe. As such, it is not surprising that our neural architecture supports interactions of cognition and emotion at a number of processing levels. Emotion has been shown to influence cognitive functions such as decision making (Bechara, Damasio, & Damasio, 2003), categorisation (Ito, Larsen, Smith, & Cacioppo, 1998), and memory (Hamann, Ely, Grafton, & Kilts, 1999). A growing body of research suggests that emotion influences information processing even at early attentional (Carretié, Hinojosa, & Mercado, 2003; Smith, Cacioppo, Larsen, & Chartrand, 2003) and perceptual (Phelps, Ling, & Carrasco, 2006) stages. While there is a substantial literature dedicated to understanding cognition–emotion interactions, much of this work has focused on visually presented emotional stimuli, leaving many open questions concerning emotion processing in the auditory modality. Nevertheless, auditory emotional signals are just as salient as their visual counterparts, perhaps even more so since they can be detected from a

greater distance and in low-lighting conditions. In fact, early in development, auditory emotional signals have been shown to take precedence over facial emotion cues in social referencing (Mumme, Fernald, & Herrera, 1996). The current study examines the recognition of emotional prosody, with the aim of investigating its time course and whether valence and/or arousal influences how quickly and accurately emotion is recognised in tone of voice.

Emotion processing in the auditory modality

While the literature on affective prosody is substantial, the majority of studies have focused either on the neural localisation of emotional prosody recognition or on its acoustic correlates rather than on potential processing differences between emotional and neutral prosody. Nevertheless, results from a small number of studies bear on the influences of valence and arousal on emotional prosody processing. For instance, Alter and colleagues (2003) presented participants with utterances spoken with happy, angry, and neutral prosody. They recorded event-related potentials (ERPs) as participants judged the prosody on a scale of 1 (*negative*) to 5 (*positive*) and found that the amplitude of the P200 component, thought to index early sensory stages of feature detection (Luck & Hillyard, 1994), was larger in response to happy compared to angry and neutral prosody, indicating enhanced processing of positive emotion. Other studies have instead suggested enhanced processing of negative prosody. For instance, Wambacq and colleagues (Wambacq, Shea-Miller, & Abubakr, 2004) compared processing of negative and neutral prosody and found that the amplitude of the P3 ERP component (thought to reflect the categorisation of anomalous, inconsistent or infrequent stimuli presented in a context of otherwise normal or frequent information) was larger in response to negative prosody. In addition, in an fMRI study, Grandjean and colleagues (2005) found greater activation in several brain areas, including the superior temporal sulcus, when participants processed angry compared to neutral prosody. However, a positive emotion condition was not included in either Wambacq's (2004) or Grandjean's (2005) study, so the data could reflect either a processing bias for negative prosody or a more general processing advantage for emotional (or high-arousal) compared to neutral input. Grandjean and colleagues (Grandjean, Sander, Lucas, Scherer, & Vuilleumier, 2008) recently reported findings consistent with the latter possibility in a study of brain-damaged patients with auditory extinction, an attentional deficit in which patients fail to detect auditory information presented to the contralesional side when auditory information is presented simultaneously to the ipsilesional side. Grandjean and colleagues (2008) found that auditory extinction was attenuated for positive and negative compared to neutral prosody, suggesting that auditory emotional content, regardless of valence, captures attention and enhances perception. In addition, the amplitude of the mismatch negativity, an indicator of acoustic change detection, has been shown to be larger to both happy and angry compared to neutral prosody, although this effect was seen in women but not men (Schirmer, Striano, & Friederici, 2005).

As the aforementioned studies demonstrate, there is conflicting evidence in the emotional prosody literature regarding effects of valence and arousal on emotional prosody processing. Some studies suggest valence effects that favour either positive or negative prosody, whereas other studies suggest more general arousal effects that favour emotional over neutral prosody. There is even evidence supporting a bias for neutral over emotional prosody. In a study by Schirmer and Kotz (2003), participants heard semantically positive, negative, and neutral words, spoken with positive, negative, or neutral prosody. Their task involved categorising stimuli as positive, negative, or neutral. In one block, they identified the emotion of the words themselves while ignoring the prosody, and in another block, they identified the emotional prosody while ignoring the semantics. Results indicated that participants were more accurate in recognising neutral prosody than happy or angry prosody.

Interestingly, this effect remained constant regardless of the semantics of the words, except when the semantics were angry, in which case the recognition accuracy of angry prosody approached the accuracy level of neutral prosody. These results suggest that emotional prosody is not always privy to enhanced processing; instead, there appear to be circumstances in which neutral prosody is favoured.

Emotion processing in the visual modality

Similar to results in the auditory modality, the effects of visual emotional content on perception and cognition tend to fall into two main categories. In one, emotional stimuli, irrespective of their valence, are privy to enhanced processing compared to neutral stimuli. That is, positive and negative emotionality have similar facilitating effects on processing. Consistent with this stance, results from a study by Zeelenberg and colleagues (Zeelenberg, Wagenmakers, & Rotteveel, 2006) demonstrated that performance in a perceptual identification task was better for positive and negative compared to neutral target words. Anderson (2005) found that the attentional blink, a phenomenon in which the detection of a second target word is suppressed when it is presented at a very short latency following the first target word, was attenuated when the second target was an emotional word, suggesting increased allocation of attention to emotional stimuli. Furthermore, this attenuation was modulated by stimulus arousal level rather than valence. Results from a study in which pictures were used to test the influence of emotion on attention (Schimmack, 2005) are also consistent with the hypothesis that enhanced processing of emotional stimuli is due to their higher arousal levels compared to emotionally neutral stimuli.

In contrast to the perspective that emotion enhances processing in a valence-independent fashion, a second category of effects has indicated valence-specific effects. Some results suggest a processing advantage for happy stimuli. For instance, faster reaction times (e.g., Kirita & Endo, 1995; Leppänen & Hietanen, 2004; Leppänen, Tenhunen, & Hietanen, 2003) and greater accuracy (Calvo, Nummenmaa, & Avero, 2008) have been demonstrated when participants categorise happy faces compared to faces expressing negative emotion. However, other valence-specific findings suggest that negative stimuli have a privileged status, termed the “negativity bias”. For instance, angry faces have been reported to be detected more efficiently than happy faces (Horstmann & Bauland, 1996; Öhman, Lundqvist, & Esteves, 2001). Faster detection of negative content has been reported using a variety of paradigms, including dot probe (Mogg & Bradley, 1999) and Stroop-like (Pratto & John, 1991) tasks, different types of stimuli, including words (Dijksterhuis & Aarts, 2003), and even in conditions of restricted awareness (Dijksterhuis & Aarts, 2003).

The current study

Despite reports of processing advantages for emotional stimuli in the visual modality, findings concerning the auditory correlates are scant. Thus, despite its theoretical importance, the extent to which the effects of emotion on perception are global versus modality specific remains unclear. The purpose of the current research was to investigate the influence of emotional content on auditory processing in order to probe the existence of emotion-related biases in the auditory domain. We focused on emotion recognition in prosodically marked speech for two reasons: (a) its importance in social interactions and (b) it serves as an auditory analogue to conventionally investigated emotional facial expressions. We specifically chose to examine the processing of happy, angry, and neutral prosody in order to tease apart the potential influences of valence and arousal level, as happiness and anger are both high in arousal but opposite in valence (Feldman Barrett & Russell, 1998).

A gating paradigm was employed to identify the time point at which participants recognise different emotions. In the typical gating study (Grosjean, 1980, 1996), spoken words are spliced into segments, or gates, of increasing duration. After hearing each gate, participants guess the word they think they heard and rate their confidence in their decision. The isolation point, defined as the length (in milliseconds) of the gate at which participants correctly identify the word and don't subsequently reverse their decision, is calculated across participants and/or items. These data are used to determine how much acoustic information is necessary for participants to accurately identify a word, with faster isolation points indicating faster word recognition. Although the gating paradigm has been employed primarily at the word level (Warren & Marslen-Wilson, 1987), it is directly applicable to the study of prosodic processing at the sentential level. For instance, Grosjean and Hirt (1996) utilised a gating paradigm to investigate how people use linguistic prosody to predict the end of a sentence. Akin to this approach, here we use the gating paradigm to investigate the time course of recognising emotional prosody. A major strength of this approach is that by controlling the amount of stimulus information provided to participants in each trial, we are able to conduct a fine-grained examination of the time course of emotional prosody recognition while eliminating any potentially confounding influence of a speed–accuracy trade-off. In sum, the gating paradigm's temporal sensitivity allows for a detailed investigation of emotionally valenced prosodic information in spoken sentences; nevertheless, we are not aware of any previous applications of the paradigm to emotion-processing research.

In the current study, participants heard utterances spoken in happy, angry, and neutral prosody, which were spliced into gates of increasing duration. Each successive gate increased in duration by 250 ms, so that Gate 1 was 250 ms long, Gate 2 was 500 ms long, Gate 3 was 750 ms long, and so on. After every gate, participants made a forced choice between the emotion categories. The logic behind this investigation is that if the negativity bias reported in the visual emotion recognition literature reflects a global and not a modality-specific advantage, then participants should be faster and/or more accurate in identifying angry compared to happy or neutral emotion in auditorily presented sentences, manifested in faster isolation points for angry prosody and greater overall accuracy once the entire sentence has been heard. Similar advantages would be expected for happy prosody if the happy bias reported with face stimuli carried over to emotional prosody. Alternatively, if either the presence or the arousal level of the emotion, rather than its valence, is critical, then angry and happy prosody should be identified similarly, with greater speed and accuracy than neutral prosody. An additional possibility is that emotionally intoned speech differs from emotional visual stimuli with regard to early processing advantages. If so, then either (a) neutral prosody would be identified more quickly and accurately than emotional prosody, consistent with Schirmer and Kotz's (2003) intriguing finding, or (b) no differences in speed or accuracy between happy, angry, and neutral conditions would be expected.

EXPERIMENT 1

Method

Participants—Fifty-one undergraduate students at UC San Diego (30 female, 21 male; mean age=21 years, $SD=2.7$) were awarded psychology course credit in exchange for their participation. Participants were screened and determined via self-report questionnaire to be monolingual native English speakers. They self-reported that they had normal hearing and no history of head trauma, psychiatric illness, or other impairment that would preclude them from participating. Data from eight participants were excluded; seven due to second language exposure before age six and one due to computer failure, leaving data from 43 (27 female, 16 male) participants to be analysed.

Stimuli—Stimuli were “Jabberwocky” sentences. Jabberwocky, named for Lewis Carroll’s (1871) poem consisting of nonsense verse, is a pseudo-language in which English grammatical structure is maintained, yet utterances are semantically incomprehensible because many of the words are nonsensical. In the present study, English sentences were transformed into Jabberwocky by keeping closed class words and verbs intact but replacing the nouns with phonologically pronounceable non-words of the same syllable length (noted in the following example in *italics*: “*The hessups ate pea-chup after the sholt*”). Sentences were recorded by an actress from UCSD’s Theatre Department who was instructed to use tone of voice to convey happy, angry, or neutral emotionality. Each sentence was recorded in all three emotion conditions. By utilising Jabberwocky sentences, we ensured that many features of spoken English were preserved; however, semantic meaning (context) could not influence processing of the emotion. Instead, emotion was solely conveyed via prosody.

Pre-test 1: Confirmation of emotionality: A separate group of 40 participants (UCSD undergraduates receiving course credit) listened to a set of 110 Jabberwocky sentences spoken in happy, angry, and neutral prosody. Participants were randomly assigned to one of four stimulus lists, each containing 110 sentences of all emotion types presented in fixed random order. They were given a test booklet and told that they would hear nonsense sentences spoken in happy, angry, or neutral tones of voice. During a 7 second period of silence in between sentences, they were to first indicate the valence of the sentence by making an “x” in the happy, angry, or neutral column. Second, because it was important based on our hypotheses to select happy and angry stimuli that were similar in arousal level, participants were asked to rate on a scale from 1 (*very weakly*) to 5 (*very intensely*) how strongly each sentence invoked the emotion they selected. (Note that for neutral items, participants rated how strongly the neutral stimuli evoked a feeling of “neutrality”).

Final stimulus set: For this study, we selected the 48 sentences that were most accurately recognised across *all three* emotion categories by participants in Pre-test 1, yielding 144 total stimuli. The mean recognition percentage for the stimuli included in the final set was 95% ($SD=8$) overall, 98% ($SD=6$) for happy stimuli, 97% for angry stimuli ($SD=6$), and 92% ($SD=10$) for neutral stimuli. Subsequently, using Cool Edit Pro software (Syntrillium Software, 2003), maximum sound amplitudes were roughly equated across all sentences.

Pre-test stimulus ratings: For the 48 sentences in the final stimulus set, ratings of perceived emotion strength, obtained in Pre-test 1, were compared across emotions. The ratings of the emotionally charged stimuli were rated similarly (for happy and angry items, $M_s=3.6$ and 3.7 and $SD_s=0.57$ and 0.59 , respectively). The neutral items received ratings that were weaker overall ($M=3.3$, $SD=0.42$) than the happy and angry items. An analysis of variance (ANOVA) revealed a main effect of emotion, $F_2(2, 46)=7.30$, $p=.001$, $\eta_p^2=.134$. As expected, pairwise comparisons indicated lower emotion strength ratings for neutral than for happy ($p=.005$) and angry sentences ($p=.002$), whereas happy and angry sentences did not differ from one another ($p=1.00$).

Stimulus duration: In the final stimulus set, the average sentence duration was 2.7 s overall [range= 1.6 s to 4.4 s], with happy=2.8 s, angry=2.9 s and neutral=2.5 s. An ANOVA revealed a main effect of emotion on duration, $F_2(2, 46)=43.56$, $p=.000$, $\eta_p^2=.481$, with neutral sentences shorter than happy and angry sentences ($p_s=.000$).¹

¹At the time of recording, it was felt that preservation of the naturalistic components of the speech samples was critical for this experiment. The differences in sentence length are considered in the analyses (see results).

Stimulus acoustics for full sentences: Based on previous research suggesting the acoustic properties that are most crucial for conveying emotion in speech (Mozziconacci, 2001; Scherer, 1986), pitch minimum, maximum, range, mean, and standard deviation, as well as mean intensity of the 48 sentences included in the final stimulus set were extracted from the sound files via Praat 4.6.12 (Boersma & Weenink, 2007), a computer program for speech analysis. These properties were then compared across emotion categories using ANOVAs. Means and standard deviations are presented in Table 1. With respect to minimum and maximum pitch, ANOVAs revealed a significant effect of emotion, $F_{2s}(2, 46)=6.91$ and 33.27 , $ps=.003$ and $.000$, $\eta_p^2=.128$ and $.414$, respectively. Pairwise comparisons demonstrated that the minimum pitch of happy utterances was higher than angry utterances ($p=.001$), but no other comparisons reached statistical significance. Happy stimuli had a higher maximum pitch than both angry and neutral stimuli (both $ps=.000$), but angry and neutral stimuli did not differ from one another ($p=1.00$). The effect of emotion was also significant for pitch range, $F_2(2, 46)=15.62$, $p=.000$, $\eta_p^2=.249$. Happy sentences had a greater pitch range than both angry and neutral sentences (both $ps=.000$), which did not differ from each other ($p=.761$). There was also a significant effect of emotion on mean pitch, $F_2(2, 46)=217.34$, $p=.000$, $\eta_p^2=.822$. The mean pitch of happy stimuli was higher than that of angry stimuli ($p=.000$), which was in turn higher than that of neutral stimuli ($p=.000$). There was a similar effect of emotion on the standard deviation of pitch, $F_2(2, 46)=91.04$, $p=.000$, $\eta_p^2=.660$, which was greater for happy than angry ($p=.000$) and neutral ($p=.000$) stimuli, and greater for angry than neutral stimuli ($p=.001$). Lastly, there was a significant effect of emotion on mean intensity, $F_2(2, 46)=37.30$, $p=.000$, $\eta_p^2=.442$, reflecting higher values for happy compared to angry and neutral sentences (both $ps=.000$), which did not differ from one another ($p=1.00$).

Preparation of stimulus gates: Once chosen, the sentences were then prepared for the gating paradigm by editing each sentence into successive 250 ms clips, or gates. In between each gate was a 5 s period of silence (see Figure 1). Thus, for every sentence, Gate 1 consisted of the first 250 ms of the sentence, Gate 2 consisted of the first 500 ms of the sentence, Gate 3 consisted of the first 750 ms of the sentence, and so on until the end of the sentence. The number of gates ranged from 7 to 18, with a mean of 11 ($SD=2$).

Stimulus acoustics for Gates 1–4: The same acoustic properties were analysed for the first (250 ms in length) through fourth (1000 ms in length) gates of each sentence in order to examine potential differences in how the acoustics of emotional prosody built over the early part of the sentences. Means and standard deviations are presented in Table 1. First, ANOVAs with Emotion (happy, angry, neutral) and Gate Number (1, 2, 3, 4) as repeated measures were carried out. The most interesting results concern potential interactions of emotion and time, which bear on the ways in which the prosodic cues conveying each emotion unfold over time by elucidating the time points at which various acoustic properties corresponding to each emotion begin diverge from one another. Next, follow-up ANOVAs for each variable were conducted to examine effects of emotion separately for each gate. Significance levels for main effects of these follow-up analyses were adjusted to $\alpha=.0125$ according to a Bonferroni procedure, and violations of the sphericity assumption were treated with a Greenhouse–Geisser correction. Results from these follow-up analyses are presented in Table 2.

An ANOVA for minimum pitch revealed main effects of both Emotion, $F_2(2, 46)=19.347$, $p=.000$, $\eta_p^2=.292$, and Gate Number, $F_2(3, 45)=40.724$, $p=.000$, $\eta_p^2=.464$, but no significant interaction of Emotion and Gate Number ($p=.782$). For maximum pitch, main effects of

Emotion and Gate Number again reached significance, $F_2(2, 46) = 50.801, p = .000, \eta_p^2 = .519$, and $F_2(3, 45) = 54.819, p = .000, \eta_p^2 = .538$, respectively. Here, the interaction was also significant, $F_2(6, 42) = 4.432, p = .003, \eta_p^2 = .086$. For pitch range, again both main effects were significant, $F_2(2, 46) = 27.280, p = .000, \eta_p^2 = .367$ for Emotion and $F_2(3, 45) = 96.576, p = .000, \eta_p^2 = .673$ for Gate Number, as was their interaction, $F_2(6, 42) = 4.715, p = .003, \eta_p^2 = .091$. Analysis of mean pitch revealed significant main effects of Emotion and Gate Number, $F_2(2, 46) = 87.857, p = .000, \eta_p^2 = .651$, and $F_2(3, 45) = 24.539, p = .000, \eta_p^2 = .343$, respectively, and their interaction, $F_2(6, 42) = 8.387, p = .000, \eta_p^2 = .151$. An ANOVA on the standard deviation of pitch demonstrated significant main effects of both Emotion, $F_2(2, 46) = 31.918, p = .000, \eta_p^2 = .404$ and Gate Number, $F_2(3, 45) = 19.804, p = .000, \eta_p^2 = .296$, qualified by a significant interaction of Emotion and Gate Number, $F_2(6, 42) = 2.370, p = .030, \eta_p^2 = .048$. Lastly, for mean intensity, significant main effects of Emotion, $F_2(2, 46) = 13.029, p = .000, \eta_p^2 = .217$, and Gate Number, $F_2(3, 45) = 97.073, p = .000, \eta_p^2 = .674$ emerged, as well as a significant interaction, $F_2(6, 42) = 8.842, p = .000, \eta_p^2 = .158$. In sum, five out of six acoustic variables exhibited significant interactions of emotion and gate number. As shown in Table 2, follow-up analyses indicated that happy prosody tended to differ from angry and neutral prosody along the chosen dimensions, similar to the results of analyses based on the full sentences, described above.

Design—In this mixed-factorial design, all participants heard sentences expressing all three emotions. However, to minimise potential repetition effects and biases, they did not hear the same sentence across all three emotions. Thus, three lists were created, in which sentences and emotions were counterbalanced. Each list consisted of 48 Jabberwocky sentences: sixteen were spoken in happy intonation, 16 in angry intonation, and 16 in neutral intonation. These sentences were presented in fixed random order, with the caveat that no more than three sentences expressing the same emotion occurred in a row.

Procedure—Participants were randomly assigned to one of the three stimulus lists, given a test booklet, and instructed that they would hear short audio clips of people speaking nonsense words in sentences. They were told that each sentence would be broken up into successively longer pieces, with the full sentence being presented as the final audio clip in that set. After hearing every clip, there would be a five-second period of silence which would serve as a cue for them to write a mark in the appropriate column in their test booklet to indicate whether the clip conveyed a happy, angry, or neutral tone, and to rate their confidence in that decision on a scale from 1 (*very unsure*) to 5 (*very sure*). A sequence of three beeps signalled the beginning of the next sentence. After receiving instructions, participants first responded to a practice sentence and were given feedback. Clips were presented through stereo headphones in an individual sound-attenuated booth.

Measures of interest—Analyses were focused on four main variables: First, we assessed participants' accuracy in recognising happy, angry, and neutral prosody, quantified as their percentage correct across items in each emotion condition. Second, to evaluate the time course of processing emotional prosody, we identified each participant's isolation point for each item (Grosjean, 1980, 1996), operationalised as the length of the gate (i.e., 250 ms, 500 ms, 750 ms, etc.) at which participants chose the appropriate emotion and did not subsequently change their decision. These lengths were then averaged for each participant across the sentences in each of the three emotion categories (happy, angry, and neutral).² Third, we examined participants' patterns of errors, as a negativity bias could be reflected in

a propensity to label neutral prosody as angry or happy prosody as neutral. Fourth, we extracted participants' confidence ratings (on a 1–5 scale) for each sentence at two time points of interest: the isolation point and the final gate (when the sentence was heard in its entirety). For both time points, we averaged the confidence ratings of each participant separately for happy, angry, and neutral sentences. Confidence at the isolation point and at the final gate were analysed as a means of testing for a potential emotion-related bias in the degree of conviction with which participants identified happy, angry, and neutral prosody. When appropriate, additional analyses were conducted (see data analysis below) to further elucidate the nature of significant effects.

Data analysis—All data were analysed using SPSS 11.5. We submitted variables of interest to repeated-measures ANOVAs. We conducted two types of ANOVAs: standard subject (F_1) analyses in which we collapsed across items, and also item (F_2) analyses in which we collapsed across subjects. Just as results from subject analyses provide information on the generalisability of an effect across subjects, item analyses provide analogous information on the generalisability of an effect across stimuli. Emotion (happy, angry, and neutral) was a within-subjects factor in F_1 and F_2 analyses, and subject analyses also included Stimulus List (1, 2, 3) and participant Gender (male, female) as between-subjects factors.³ Pairwise comparisons in conjunction with each variable of interest used a Bonferroni adjustment to preserve a familywise $\alpha=.05$.

Results

Percent correct—In order to investigate whether emotional prosody would be recognised more easily than neutral prosody, participants' recognition accuracy was compared for happy, angry, and neutral sentences. Percentage correct was computed for all 43 participants ($M=85\%$, $SD=9.0$). Two participants were excluded as they scored more than 2 standard deviations below the overall mean. Therefore, 41 participants remained for final analysis.

Computed separately for each emotion, participants achieved an average percentage correct of 81% ($SD=12.6$) for happy, 87% ($SD=11.1$) for angry, and 92% ($SD=10.1$) for neutral stimuli. With percent correct as the dependent measure, repeated-measures ANOVAs were performed with Emotion as a within-subjects factor and Stimulus List and participant Gender as between-subjects factors in the subject analysis. There were no main effects of Stimulus List ($p=.89$) or Gender ($p=.58$). There was a main effect of Emotion found in both the subject and item analyses, $F_1(2, 34) = 7.87$, $p=.001$, $\eta_p^2=.184$; $F_2(2, 46)=5.87$, $p=.004$, $\eta_p^2=.111$. Pairwise comparisons indicated that participants were more accurate when identifying neutral as compared to happy or angry prosody ($ps=.001$ and $.045$, respectively). There was no difference in accuracy between happy and angry sentences ($p=.72$; see Figure 2).

Isolation point—By definition, the isolation point is the length (in milliseconds) of the gate at which participants first correctly identify the emotion conveyed without changing their mind at subsequent gates. Therefore, the isolation point for a given subject for a given sentence is a multiple of 250 (e.g., 250 ms, 500 ms, 750 ms, etc.). To test whether emotional prosody was recognised faster than neutral prosody, we first identified each participant's

²The term "isolation point" was borrowed from the spoken word recognition literature, in the context of which the gating paradigm was initially developed. It is expected that there will be considerable individual differences in participants' identification of emotional prosody. Those individual differences, coupled with the longer stimulus durations of full sentences compared to single words, will likely lead to increased variability in isolation points in the current experiments compared to word-recognition studies. Nevertheless, the original terminology is retained for methodological consistency.

³Gender differences were not the focus in the current study; however, because they have been observed in some previous studies of emotional prosody processing (e.g., Schirmer et al., 2005), we included gender as a factor in data analyses.

isolation points for all 48 sentences they heard. We then averaged each participant's isolation points across happy, angry, and neutral categories. These numbers were at least 250 ms but because they were averages, they were not limited to 250 ms increments and instead usually fell somewhere in-between gate lengths (e.g., 600 ms).

Overall, participants isolated the emotion conveyed after hearing an average of 802 ms ($SD=258$) for happy, 723 ms ($SD=233$) for angry, and 444 ms ($SD=192$) for neutral items. A repeated-measures ANOVA revealed a significant effect of emotion on isolation point by participants, $F_1(2, 34)=24.67$, $p=.000$, $\eta_p^2=.413$, and by items, $F_2(2, 46)=13.81$, $p=.000$, $\eta_p^2=.227$ (see Figure 3). Pairwise comparisons indicated faster correct identification of neutral compared to both happy and angry prosody ($ps=.000$ and $.001$, respectively, for subject and item analyses). Angry prosody was also identified faster than happy prosody ($p=.04$), although this contrast was not significant in the item analysis ($p=.71$). There were no effects of Stimulus List or Gender in the subject analysis.

Given the discrepancy in stimulus duration between emotions, perhaps the faster identification of neutral sentences resulted from their shorter length, with quicker isolation points reflecting an advantage of having heard more information by a given time point. To test this, we computed a ratio of isolation point to stimulus duration for each item and conducted an ANOVA with emotion as a repeated measure. A significant effect would rule out the possibility that the faster isolation points for neutral prosody were due to the shorter duration of those stimuli, as the ratio corrected for differences in length between emotional conditions. Indeed, the result remained significant, $F_2(2, 46)=11.82$, $p=.000$, $\eta_p^2=.201$, and pairwise comparisons indicated that the ratio was smaller for neutral than for happy and angry items ($ps=.000$ and $.01$, respectively). Ratios for happy and angry items did not differ ($p=.29$). Therefore, the faster identification of neutral prosody cannot be accounted for by differences in sentence duration.

Neutral by default or a distinct processing advantage?: To probe the faster isolation points for neutral sentences, we examined participants' response patterns for the first three gates (Table 3). In particular, we investigated whether participants selected neutral at the first gate more frequently than happy or angry, either due to a perceptual or response bias, or perhaps because the sentences may not have conveyed a definite emotion in the first 250 ms. If participants disproportionately chose to label sentences as neutral following the initial gate, then it is possible that the effect of emotion on isolation point can be explained by a propensity to select neutral by default, rather than a processing advantage for neutral prosody. To examine this, the number of times participants incorrectly selected "neutral" at the first 250 ms gate was compared to chance performance, where chance was set at 10.67, since participants were presented with 32 non-neutral items and three possible emotional labels. Incorrect selection of neutral was found to occur at chance level, $t_1(40)=-0.036$, $p=.97$ (2-tailed), $d=0.01$, suggesting that participants were not using neutral as an initial default choice, because had that been the case, neutral would have been selected more frequently than chance occurrence. Thus, the results cannot be accounted for by a default tendency to label segments as neutral.

Error types—In order to investigate response patterns for the final gate and to test for a bias for emotional over neutral prosody, participants' errors were quantified in two ways: (1) incorrect labelling of an emotion; and (2) directionality of biases (see Table 4).

Incorrect labelling: This measure was the number of errors that constituted incorrectly labelling an item as happy, angry, or neutral, ignoring which of the remaining emotions was incorrectly chosen. A negativity bias in this measure would be reflected as a propensity for

participants to mislabel prosody as angry more than as happy or neutral, whereas a general emotion bias would be reflected as a propensity to mislabel prosody as happy or angry more than as neutral. A repeated-measures ANOVA revealed a significant effect of error type by subjects, $F_1(2, 34)=36.13$, $p=.000$, $\eta_p^2=.508$, and items, $F_2(2, 46)=28.13$, $p=.000$, $\eta_p^2=.374$. However, pairwise comparisons indicated that this effect did not reflect either a negativity bias or an emotion bias. Instead, sentences were incorrectly labelled as neutral more frequently than either happy or angry (both $ps=.000$), suggesting a bias for neutral prosody.

Directionality of bias: As stated earlier, a negative bias could be reflected by labelling neutral prosody as angry or happy prosody as neutral. Alternatively, a positive bias could be reflected in a propensity to label angry prosody as neutral or neutral prosody as happy. Each type of bias was operationalised as the sum of its component error types. A paired-samples t -test to compare these two types of biases, although not significant, suggested a trend toward a negative bias by subjects, $t_1(40)=1.65$, $p=.055$ (1-tailed), $d=0.385$, and items, $t_2(47)=1.60$, $p=.06$ (1-tailed), $d=0.304$.

Confidence ratings—We next tested for negativity and emotion biases in participants' confidence ratings. Higher confidence ratings for angry stimuli compared to happy and neutral stimuli would indicate a negativity bias. Higher confidence ratings for angry and happy compared to neutral stimuli would indicate a more general bias for emotional prosody. Confidence ratings for the first five gates and the last gate, in addition to the average accuracy at each of those gates, are presented in Table 5. Two analyses were carried out: First, in order to validate the confidence ratings measure, we averaged ratings at the final gate for correct and incorrect items (collapsed across emotions) and compared these two conditions using a paired-samples t -test. Mean ratings were 4.30 ($SD=0.48$) and 3.18 ($SD=0.83$) for correct and incorrect items, respectively, and participants were more confident on items they labelled correctly, $t_1(38)=11.86$, $p=.000$ (2-tailed), $d=1.632$.

The second analysis had two aims: First, to further validate the confidence ratings measure, we tested, for correct items only, whether participants were more confident of the emotion conveyed at the final gate (after hearing the entire sentence) than at the isolation point. Second, we tested for emotion biases by analysing whether confidence differed between emotions (see Table 6). Therefore, a repeated-measures ANOVA with Emotion and Time (isolation point, final gate) as within-subjects factors, and Stimulus List and Gender as between-subjects factors, was conducted. Results revealed a main effect of Time, as predicted, with participants reporting greater confidence in their emotion classifications at the final gate (after hearing the whole sentence) than at the isolation point, $F_1(1, 35)=266.54$, $p=.000$, $\eta_p^2=.884$, but no effects or interactions involving Emotion, Stimulus List, or Gender (all $ps>.15$). An ANOVA by items confirmed the main effect of Time, $F_2(1, 47)=2515.29$, $p=.000$, $\eta_p^2=.982$; however, here the main effect was qualified by a significant interaction of Emotion and Time, $F_2(2, 46)=4.54$, $p=.01$, $\eta_p^2=.088$. Follow-up ANOVAs were conducted to elucidate the nature of this interaction. Although there was no effect of Emotion on confidence ratings at the isolation point ($p=.58$), the effect was significant at the final gate, $F_2(2, 46)=4.37$, $p=.02$, $\eta_p^2=.085$. Pairwise comparisons indicated that at the final gate, confidence was greater for neutral than for happy sentences ($p=.02$), and for neutral than angry sentences ($p=.05$).

Discussion

Contrary to previous studies involving visual stimuli, which indicate a privileged role for emotional content in information processing, the present results suggest an advantage for

neutral content in the processing of emotional prosody. By utilising a gating paradigm to examine the time course of recognising happy, angry, and neutral prosody in sentence form, we first demonstrated that participants were most accurate at identifying neutral rather than positive or negative emotional prosody. Neutral prosody was identified more rapidly than either happy or angry prosody, even after correcting for the difference in stimulus duration between emotions. This finding was not due to a propensity for participants to select neutral at the first segment and persist with that label unless strongly convinced otherwise, as the number of times neutral was selected following the first segment did not differ from what would be expected due to chance.

Close examination of error patterns revealed differences between emotions. Participants were more likely to incorrectly label the prosody of a sentence as “neutral” than as “happy” or “angry”, suggesting a neutral bias in prosodic processing; however, the nature of this bias is unclear: Although the error patterns raise the possibility of a response bias, the aforementioned result that participants did not appear to identify prosody as neutral by default suggests that a response bias could not fully account for the data. Interestingly, more errors reflected a bias in the negative rather than the positive direction, which might be expected to result from a processing advantage for negative stimuli, although this comparison did not reach statistical significance. A potential negativity bias was also suggested by participants’ faster identification of angry compared to happy prosody (see Figure 3). However, this finding was overshadowed by the fact that the fastest isolation points occurred in the neutral condition.

Analyses on confidence ratings at the point of recognition and following the final gate demonstrated no differences in confidence between emotions at either time point in the subject analysis; however, the item analysis revealed that confidence at the final gate was significantly greater for neutral compared to happy sentences and marginally greater for neutral compared to angry sentences. This finding suggests an advantage for neutral intonation; nevertheless, it should be interpreted with caution since it was apparent only when collapsing across subjects and using items as the dependent measure.

One possible explanation for the findings in this experiment other than a true processing bias or advantage for identifying neutral prosody is task-related artefact. Here, participants had two emotion categories to decide between but only one neutral, or non-emotional, category. Perhaps the effects in Experiment 1 reflect greater ease in deciding whether or not a stimulus is emotional compared to deciding which particular emotion it conveys. It seems unlikely that this alternative could account for the results in Experiment 1, as analyses of error patterns revealed that participants were more likely to confuse both happy and angry prosody with neutral prosody than with each other. In addition, participants did not appear to label prosody as neutral by default, and the bias remained even after correcting for sentence duration. Nevertheless, research on facial emotion recognition has demonstrated that the categories of emotional expressions presented to participants influence their judgements of those expressions, highlighting the role of context in studies of emotion recognition (Tanaka-Matsumi, Attivissimo, Nelson, & D’Urso, 1995). Experiment 2 was designed to test whether a neutral bias would persist if participants were offered only two response categories.

EXPERIMENT 2

Results from Experiment 1 revealed greater accuracy and speed of recognition for neutral as compared to happy and angry prosody. However, the extent to which the findings were influenced by the number and nature of the response categories is unknown. In Experiment 1, participants had to choose between two emotional categories (happy and angry) and one

non-emotional category (neutral). One way to conceptualise the task is that in the emotional conditions, two decisions were required. Participants had to decide first, whether the stimulus sentence conveyed emotion or not, and second, if it did, whether that emotion was positive or negative. However, for neutral stimuli, only one decision was necessary, since once the participant decided the sentence did not convey emotion, no further decision was necessary. One possible explanation for the advantage we observed for processing neutral over emotional stimuli, then, could be that the task was easier and faster in the neutral than in the emotional conditions. Since analyses from Experiment 1 do not allow for us to rule out this possible task-related artefact, Experiment 2 reduced the number of emotion categories the participants received: they were asked to correctly recognise prosody that fell into one of only two emotion categories: either happy *or* angry, and neutral. If the neutral bias evinced in Experiment 1 reflected artefact due to the number of stimulus categories, then a neutral bias would not be expected when participants only chose between two categories of prosody. However, if Experiment 1 revealed a true neutral bias, we would expect the results of Experiment 2 to replicate those in the previous experiment, with more accurate and faster identification of neutral prosody than emotional prosody.

Method

Participants—Twenty-four undergraduate students at UC San Diego (15 female, 9 male; mean age=22 years, $SD=2.3$) were awarded psychology course credit in exchange for their participation. As in Experiment 1, participants were monolingual native English speakers, with no history of head trauma or other impairment that would preclude participation.

Stimuli—Stimuli were the same as in Experiment 1: Forty-eight Jabberwocky sentences, spoken in happy, angry, and neutral prosody, that were spliced into successive clips increasing in 250 ms increments. (See Experiment 1 for additional details.)

Design—Stimuli were divided across four lists, in which sentences and emotions were counterbalanced. Similar to Experiment 1, each list consisted of 48 Jabberwocky sentences, presented in fixed random order. Participants assigned to stimulus Lists 1 and 2 were presented with angry and neutral prosody, and participants assigned to stimulus Lists 3 and 4 were presented with happy and neutral prosody.

Procedure—Participants were randomly assigned to one of the four stimulus lists, seated in front of a computer, and instructed that they would hear short audio clips of people speaking nonsense words in sentences. As in Experiment 1, instructions indicated that each sentence was broken up into successively longer pieces, with the full sentence presented as the final audio clip in that set. A five-second period of silence following each segment served as a response prompt. After receiving instructions, participants responded to and were given feedback on a practice sentence. After hearing each segment, participants responded by pressing the "<" key on the keyboard with their dominant hand if they thought the emotion expressed was angry (stimulus Lists 1 and 2) or happy (stimulus Lists 3 and 4). All participants pressed the "?" key with their dominant hand to indicate neutral prosody. Because the neutral bias observed in the previous experiment was most apparent in accuracy and isolation points, confidence ratings were not obtained here. Sound clips were presented through stereo headphones.

Measures of interest—The two measures of interest in Experiment 2 were the accuracy and speed (indicated by the isolation points) of identifying emotional versus neutral prosody. As in Experiment 1, accuracy was determined by whether the participant labelled the emotion correctly at the final gate (i.e., after having heard each sentence in its entirety). The

isolation point was determined for each participant for every item, and then averaged for items within each emotion category and compared across emotions.

Data analysis—Data were analysed using SPSS 11.5. Matched-pairs *t*-tests were carried out for both variables of interest. For participants assigned to stimulus List 1 or 2, accuracy and isolation points were paired across angry and neutral prosody to assess potential differences between the two emotion conditions. Likewise, for participants assigned to stimulus List 3 or 4, accuracy and isolation points for identifying happy versus neutral prosody were compared. *T*-tests were one-tailed due to our a priori hypothesis that the direction of any observed effects would favour neutral prosody.

Results

Percent correct—Percentage correct was computed for all 24 participants ($M=88$, $SD=8.3$). One participant's data were excluded for scoring more than 2 standard deviations below the overall mean, leaving 23 participants in the final analysis. Of those 23 participants, 11 heard angry and neutral prosody, and 12 heard happy and neutral prosody.

The first analysis was aimed at determining whether neutral prosody was identified with greater accuracy than emotional prosody. To that end data from all 23 participants were considered, regardless of whether the emotional prosody they were exposed to was happy or angry. Mean accuracy for recognising emotional intonation (collapsed across happy and angry stimuli) was 79.35% ($SD=14.50$), whereas mean accuracy for identifying neutral intonation was 96.01% ($SD=6.0$). A matched-pairs *t*-test revealed that this difference reached statistical significance, $t_1(22)=5.476$, $p=.000$, $d=1.50$. Next, accuracy data from participants assigned to stimulus Lists 1 and 2 were analysed separately from data of participants assigned to stimulus Lists 3 and 4 to examine angry and happy prosody individually. Participants assigned to Lists 1 and 2 correctly labelled 82.58% ($SD=13.02$) of angry stimuli, on average. Their corresponding accuracy for neutral stimuli was 98.12% ($SD=4.32$), which was significantly more accurate than that of angry prosody, $t_1(10)=4.479$, $p=.001$, $d=1.60$. The responses of participants assigned to Lists 3 and 4 followed a similar pattern: mean accuracy was 76.39% ($SD=15.62$) for happy sentences and 94.10% ($SD=6.76$) for neutral prosody. This difference was also statistically significant, $t_1(11)=3.522$, $p=.003$, $d=1.47$.

Isolation point—As with percent correct, data from all participants were first compiled and considered in a single analysis that compared responses to emotional (happy or angry) versus neutral sentences. The mean isolation point was 763.41 ms ($SD=520.45$) for emotional prosody and 403.93 ms ($SD=199.78$) for neutral prosody. A matched-pairs *t*-test confirmed that participants were faster at recognising neutral prosody than emotional prosody, $t_1(22)=3.751$, $p=.001$, $d=0.91$. Next, we analysed the data separately for participants who heard angry and neutral prosody (stimulus Lists 1 and 2) and those who heard happy and neutral prosody (stimulus Lists 3 and 4) in order to see whether the same patterns were evident in both cases. Participants assigned to Lists 1 and 2 isolated angry prosody after hearing an average of 841.24 ms ($SD=594.18$) and neutral prosody after hearing an average of 437.94 ms ($SD=262.40$). A matched-pairs *t*-test revealed that the isolation points were faster for neutral compared to angry prosody, $t_1(10)=3.061$, $p=.006$, $d=0.88$. Mean isolation points for participants assigned to Lists 3 and 4 were 692.07 ms ($SD=457.36$) for happy prosody and 372.75 ms ($SD=122.52$) for neutral prosody. Here again, the two emotion conditions significantly differed from one another, $t_1(11)=2.234$, $p=.024$, $d=0.95$, with faster isolation points for neutral prosody.

Discussion

Experiment 1 demonstrated that participants were more accurate and needed less auditory information to recognise neutral prosody compared to happy and angry prosody. After ruling out that these effects were due to stimulus artefact or a default tendency to select neutral, it remained possible that the neutral bias was an artefact of the stimulus categories: Specifically, we considered the possibility that perhaps it is easier and less time-consuming to make an initial decision regarding whether a stimulus is emotional or neutral but in cases where the stimulus is deemed emotional it may be more difficult to identify a specific emotion. If that were true, then neutral prosody would be identified more accurately and quickly than either happy or angry prosody, not because of a neutral bias but rather because there were two emotional stimulus categories and only one neutral category. To test this hypothesis, we conducted a second experiment in which participants were only exposed to neutral prosody and prosody expressing one emotion. The results of Experiment 2 replicate the most compelling findings from Experiment 1: participants' accuracy was greater and their isolation points were faster for the recognition of neutral prosody than happy or angry prosody. Importantly, these results demonstrate this neutral bias even when participants were presented with neutral stimuli and stimuli from only one emotion category. Therefore, the bias was not due to there being two emotional categories to decide between but only one neutral category.

GENERAL DISCUSSION

The current research was aimed at determining whether either an emotion-related or a negativity bias, both well-documented phenomena in the visual emotion recognition literature, would carry over to auditory emotion, and specifically to emotional prosody. In order to test this, we presented participants with semantically devoid Jabberwocky sentences, spoken in happy, angry, and neutral intonation. We utilised a gating paradigm, wherein the sentences were spliced into segments increasing in length, and participants indicated following every segment whether the emotion conveyed was happy, angry, or neutral. This paradigm is uniquely suited for such investigations as its temporal sensitivity facilitates examination of the time course of processing at the sentential level (Grosjean, 1980). We hypothesised that a bias for emotional or high-arousal content in prosodic processing would manifest as greater accuracy and/or less stimulus input needed to correctly identify happy and angry compared to neutral intonation. Alternatively, a bias specific to negative content would reveal greater accuracy and/or faster isolation points for angry compared to both happy and neutral stimuli, whereas a bias specific to happy content would reveal greater accuracy and/or faster isolation points for happy compared to angry or neutral stimuli. Nevertheless, our data were not supportive of either hypothesis. Instead, our results indicate that accuracy, speed of processing, and error patterns favoured neutral prosody over both happy and angry prosody.

Although findings concerning emotion-related biases in auditory emotion processing are inconsistent, our results were unexpected given the substantial literature demonstrating emotion biases when stimuli are presented visually and also because intuitively, emotional stimuli appear to be more salient than their neutral counterparts. Nevertheless, as mentioned previously, a similar "neutral bias" emerged in a study examining gender differences in emotional prosody processing (Schirmer & Kotz, 2003): In a Stroop-like task in which participants labelled the prosody of semantically positive, negative, and neutral words that were spoken with positive, negative, or neutral intonation, accuracy was greatest for detecting neutral prosody regardless of the semantics of the words. However, because this neutral prosody effect was not central to the aims of the study, the authors did not interpret the finding.

Although a neutral bias is somewhat surprising, a number of potential factors might be able to account for it. One possibility is that emotional prosody competes with other aspects of spoken language processing for limited cognitive resources. If that were the case, then the presence of emotion in the intonation patterns might lead to slower and more error-prone processing compared to neutral prosody. Another possibility relates to task relevance: in the current study, participants were explicitly instructed to identify the emotion in each segment of speech that they heard. However, perhaps better and faster detection of neutral prosody is specific to conditions where emotionality is integral to the task, whereas a bias for emotional input would emerge in incidental or task-irrelevant conditions. A previous ERP study demonstrated that negative prosody was processed more rapidly under task-irrelevant compared to task-relevant conditions but that the processing of neutral prosody did not differ between the two conditions (Wambacq et al., 2004). In addition, in a study using functional MRI, Grandjean and colleagues (2005) reported greater activation in the middle superior temporal sulcus in response to angry compared to neutral prosody in two tasks where the emotional content of the stimuli was irrelevant. These studies demonstrate that enhanced processing of emotional prosody is found when the emotional content is irrelevant to the task. The results of the present study do not rule out the possibility that processing biases for emotional prosody may be modulated by task relevance, such that the presence of emotion is more likely to enhance processing when participants' attention is not focused on the emotion in the voice per se but on other stimulus properties. Indeed, processing biases favouring emotional stimuli that are outside of the immediate focus of attention would serve an important adaptive function by drawing attention to the emotional content for further processing. Additional research examining the effects of task relevance and valence (and their interaction) on the processing of emotional prosody is warranted, especially because it has been suggested that in the visual modality, a negativity bias results when tasks tap rapid allocation of attention, whereas a positive emotion advantage results when more detailed analysis of the emotional meaning of stimuli is required (Leppänen et al., 2003).

In addition to the possibility that task-irrelevant emotional prosody or different tasks might furnish results unlike the faster and more accurate identification of neutral stimuli in the current experiment, it is possible that a crucial contributor to our results is the dynamic nature of our stimuli. The stimuli in the current experiments were designed to allow the prosodic cues to build over time, as is the case in everyday social interactions. Conversely, studies of visual emotion processing typically employ static images of emotional facial expressions. This difference could contribute to the discrepancy in findings between the current work and previous studies in the visual modality. Future research comparing the processing of static and dynamic emotion in both visual and auditory modalities would help to clarify this point. Furthermore, it is possible that different findings would emerge for classes of auditory stimuli other than prosody. Perhaps non-speech emotional vocalisations such as screams or laughter would elicit better identification than their neutral counterparts. The focus of the present work was prosody, as prosody is arguably the most ubiquitous type of auditory cue to others' emotional states. Nevertheless, future studies which test for emotion-related processing advantages for non-speech stimuli are necessary to more fully determine the extent to which emotional versus neutral stimuli are differentially processed in the auditory modality. Furthermore, studies that test auditory processing of a wide range of emotions would help to clarify the degree of specificity in the neutral bias observed in the current experiments. Given the functional similarities argued in the literature for fearful and angry emotional expression (e.g., Adolphs et al., 1999; see Öhman & Mineka, 2001, for opposing view), we chose anger as the negative emotion because it is similar in arousal, but opposite in valence, to happiness, whereas fear might differ from happiness along both dimensions. However, it is possible that the observed pattern of results would differ had we chosen fearful rather than angry prosody for our negative emotion condition.

Although our results suggest that neutral prosody is identified more accurately and rapidly than emotional prosody and therefore neither a negativity bias nor a more general emotion bias extends to this domain, at present it is unclear whether the “neutral bias” reflects a bias in perception, attention, or is a by-product of a language-processing parameter. We are currently conducting a series of experiments using ERPs to explore these alternatives.

Acknowledgments

This work was supported in part by an NSF Graduate Research Fellowship to the first author and NIH grants (DC00494 and DC03885) to the last author. Lauren Cornew is now in the Department of Radiology at Children’s Hospital of Philadelphia.

We thank Jessica Belisle, Teresa Lee, and Chris Lonner for assistance with data collection, Sarah Callahan for assistance during manuscript preparation, and three anonymous reviewers for their helpful suggestions.

References

- Adolphs R, Tranel D, Hamann S, Young AW, Calder AJ, Phelps EA, et al. Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsychologia*. 1999; 37:1111–1117. [PubMed: 10509833]
- Alter K, Rank E, Kotz SA, Toepel U, Besson M, Schirmer A, et al. Affective encoding in the speech signal and in event-related brain potentials. *Speech Communication*. 2003; 40:61–70.
- Anderson AK. Affective influences on the attentional dynamics supporting awareness. *Journal of Experimental Psychology: General*. 2005; 134:258–281. [PubMed: 15869349]
- Bechara A, Damasio H, Damasio AR. Role of the amygdala in decision-making. *Annals of the New York Academy of Sciences*. 2003; 985:356–369. [PubMed: 12724171]
- Boersma, P.; Weenink, D. Praat (Version 4.6.12) [computer software]. 2007. (Retrieved from <http://www.fon.hum.uva.nl/praat/>)
- Calvo MG, Nummenmaa L, Averro P. Visual search of emotional faces: Eye-movement assessment of component processes. *Experimental Psychology*. 2008; 55:359–370. [PubMed: 19130762]
- Carretié L, Hinojosa JA, Mercado F. Cerebral patterns of attentional habituation to emotional visual stimuli. *Psychophysiology*. 2003; 40:381–288. [PubMed: 12946112]
- Carroll, L. *Through the looking glass, and what Alice found there*. London: Macmillan; 1871.
- Dijksterhuis A, Aarts H. On wildebeests and humans: The preferential detection of negative stimuli. *Psychological Science*. 2003; 14:14–18. [PubMed: 12564748]
- Feldman Barrett L, Russell JA. Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*. 1998; 74:967–984.
- Grandjean D, Sander D, Lucas N, Scherer KR, Vuilleumier P. Effects of emotional prosody on auditory extinction for voices in patients with spatial neglect. *Neuropsychologia*. 2008; 46:487–496. [PubMed: 17945316]
- Grandjean D, Sander D, Pourtois G, Schwartz S, Seghier M, Scherer KR, et al. The voices of wrath: Brain responses to angry prosody in meaningless speech. *Nature Neuroscience*. 2005; 8:145–146.
- Grosjean F. Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*. 1980; 28:267–283. [PubMed: 7465310]
- Grosjean F. Gating. *Language and Cognitive Processes*. 1996; 11:597–604.
- Grosjean F, Hirt C. Using prosody to predict the end of sentences in English and French: Normal and brain-damaged subjects. *Language and Cognitive Processes*. 1996; 11:107–134.
- Hamann SB, Ely TD, Grafton ST, Kilts CD. Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nature Neuroscience*. 1999; 2:289–293.
- Horstmann G, Bauland A. Search asymmetries with real faces: Testing the anger-superiority effect. *Emotion*. 1996; 6:193–207. [PubMed: 16768552]
- Ito TA, Larsen JT, Smith NK, Cacioppo JT. Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology*. 1998; 75:887–900. [PubMed: 9825526]

- Kirita T, Endo M. Happy face advantage in recognizing facial expressions. *Acta Psychologica*. 1995; 89:149–163.
- Leppänen JM, Hietanen JK. Positive facial expressions are recognized faster than negative facial expressions, but why? *Psychological Research*. 2004; 69:22–29. [PubMed: 14648224]
- Leppänen JM, Tenhunen M, Hietanen JK. Faster choice-reaction times to positive than to negative facial expressions: The role of cognitive and motor processes. *Journal of Psychophysiology*. 2003; 17:113–123.
- Luck SJ, Hillyard SA. Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*. 1994; 31:291–308. [PubMed: 8008793]
- Mogg K, Bradley BP. Orienting of attention to threatening facial expressions presented under conditions of restricted awareness. *Cognition and Emotion*. 1999; 13:713–740.
- Mozziconacci SJL. Modeling emotion and attitude in speech by means of perceptually based parameter values. *User Modeling and User-Adapted Interaction*. 2001; 11:297–326.
- Mumme DL, Fernald A, Herrera C. Infants' responses to facial and vocal emotional signals in a social referencing paradigm. *Child Development*. 1996; 67:3219–3237. [PubMed: 9071778]
- Öhman A, Lundqvist D, Esteves F. The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology*. 2001; 80:381–396. [PubMed: 11300573]
- Öhman A, Mineka S. Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*. 2001; 108:483–522. [PubMed: 11488376]
- Phelps EA, Ling S, Carrasco M. Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science*. 2006; 17:292–299. [PubMed: 16623685]
- Pratto F, John OP. Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*. 1991; 61:380–391. [PubMed: 1941510]
- Scherer KR. Vocal affect expression: A review and a model for future research. *Psychological Bulletin*. 1986; 99:143–165. [PubMed: 3515381]
- Schimmack U. Attentional interference effects of emotional pictures: Threat, negativity, or arousal? *Emotion*. 2005; 5:55–66. [PubMed: 15755219]
- Schirmer A, Kotz SA. ERP evidence for a sex-specific Stroop effect in emotional speech. *Journal of Cognitive Neuroscience*. 2003; 15:1135–1148. [PubMed: 14709232]
- Schirmer A, Striano T, Friederici AD. Sex differences in the preattentive processing of vocal emotional expressions. *NeuroReport*. 2005; 16:635–639. [PubMed: 15812323]
- Smith NK, Cacioppo JT, Larsen JT, Chartrand TL. May I have your attention, please: Electrocortical responses to positive and negative stimuli. *Neuropsychologia*. 2003; 41:171–183. [PubMed: 12459215]
- Syntrellium Software. Cool Edit Pro (Version 2.1) [computer software]. Phoenix, Arizona: 2003.
- Tanaka-Matsumi J, Attivissimo D, Nelson S, D'Urso T. Context effects on the judgment of basic emotions in the face. *Motivation and Emotion*. 1995; 19:139–155.
- Wambacq IJA, Shea-Miller KJ, Abubakr A. Non-voluntary and voluntary processing of emotional prosody: An event-related potentials study. *NeuroReport*. 2004; 15:555–559. [PubMed: 15094522]
- Warren P, Marslen-Wilson W. Continuous uptake of acoustic cues in spoken word recognition. *Perception & Psychophysics*. 1987; 41:262–275. [PubMed: 3575084]
- Zeelenberg R, Wagenmakers EJ, Rotteveel M. The impact of emotion on perception: Bias or enhanced processing? *Psychological Science*. 2006; 17:287–291. [PubMed: 16623684]

The hessups ate pea chup after the sholt.

Figure 1.
Schematic of a spliced sentence with 250 ms successive clips.

Figure 2. Mean percentage correct for happy, angry, and neutral sentences. Note: *Main effect, $p=.001$; **Neutral>Happy, $p=.001$; Neutral>Angry, $p<.05$.

Figure 3. Mean isolation point (in milliseconds) for happy, angry, and neutral sentences, with lower numbers indicating faster identification of a particular emotion. Note: *Neutral<Happy, $p < .001$; **Neutral<Angry, $p = .001$; ***Angry<Happy, $p < .05$.

Table 1

Means (SD in parentheses) of acoustic properties for Gates 1–4 and overall (full sentences)

	Happy	Angry	Neutral
<i>Minimum pitch (Hz)</i>	Gate 1: 225.4 (45.5)	Gate 1: 193.0 (28.2)	Gate 1: 193.1 (21.2)
	Gate 2: 215.2 (35.6)	Gate 2: 189.4 (26.0)	Gate 2: 184.6 (13.8)
	Gate 3: 197.6 (38.9)	Gate 3: 172.7 (31.9)	Gate 3: 171.6 (24.8)
	Gate 4: 195.7 (40.2)	Gate 4: 169.5 (30.0)	Gate 4: 166.1 (27.8)
	Overall: 151.2 (35.9)	Overall: 130.5 (22.1)	Overall: 137.1 (24.7)
<i>Maximum pitch (Hz)</i>	Gate 1: 348.8 (73.4)	Gate 1: 294.8 (71.5)	Gate 1: 268.1 (47.8)
	Gate 2: 413.5 (67.0)	Gate 2: 325.7 (62.5)	Gate 2: 295.2 (47.5)
	Gate 3: 417.4 (67.6)	Gate 3: 326.5 (59.5)	Gate 3: 296.7 (49.8)
	Gate 4: 419.9 (68.0)	Gate 4: 341.9 (65.8)	Gate 4: 316.3 (73.3)
	Overall: 458.1 (57.4)	Overall: 369.1 (64.9)	Overall: 356.3 (91.0)
<i>Pitch range (Hz)</i>	Gate 1: 123.4 (67.6)	Gate 1: 101.8 (66.6)	Gate 1: 75.0 (42.6)
	Gate 2: 198.4 (68.2)	Gate 2: 136.3 (61.0)	Gate 2: 110.6 (45.9)
	Gate 3: 219.8 (74.1)	Gate 3: 153.8 (59.3)	Gate 3: 125.1 (53.3)
	Gate 4: 224.3 (76.4)	Gate 4: 172.4 (68.6)	Gate 4: 150.2 (77.9)
	Overall: 306.9 (72.2)	Overall: 238.6 (72.3)	Overall: 219.2 (102.3)
<i>Mean pitch (Hz)</i>	Gate 1: 277.1 (52.5)	Gate 1: 241.6 (41.3)	Gate 1: 232.0 (28.4)
	Gate 2: 322.1 (41.4)	Gate 2: 259.8 (32.4)	Gate 2: 240.4 (18.9)
	Gate 3: 303.4 (36.5)	Gate 3: 248.3 (30.3)	Gate 3: 227.3 (15.7)
	Gate 4: 305.7 (46.5)	Gate 4: 247.5 (29.2)	Gate 4: 224.3 (14.3)
	Overall: 289.6 (23.1)	Overall: 233.6 (19.3)	Overall: 211.3 (9.5)
<i>Pitch SD (Hz)</i>	Gate 1: 44.3 (27.2)	Gate 1: 32.8 (23.3)	Gate 1: 28.7 (16.3)
	Gate 2: 58.6 (18.2)	Gate 2: 38.8 (18.4)	Gate 2: 35.3 (12.3)
	Gate 3: 66.7 (39.0)	Gate 3: 41.6 (15.7)	Gate 3: 37.4 (12.4)
	Gate 4: 59.8 (15.6)	Gate 4: 42.5 (15.4)	Gate 4: 38.4 (12.7)
	Overall: 72.9 (16.4)	Overall: 47.7 (12.7)	Overall: 38.7 (9.6)
<i>Mean intensity (dB)</i>	Gate 1: 71.6 (4.6)	Gate 1: 70.6 (4.8)	Gate 1: 73.7 (4.5)
	Gate 2: 77.8 (1.9)	Gate 2: 75.5 (2.6)	Gate 2: 76.4 (1.6)
	Gate 3: 76.9 (1.7)	Gate 3: 74.4 (2.4)	Gate 3: 75.4 (1.5)
	Gate 4: 76.7 (1.6)	Gate 4: 74.6 (2.3)	Gate 4: 75.1 (1.5)
	Overall: 75.4 (1.2)	Overall: 73.4 (1.6)	Overall: 73.3 (1.3)

Table 2

Results of ANOVAs examining the effects of emotion on stimulus acoustic properties for Gates 1–4

	Gate 1	Gate 2	Gate 3	Gate 4
<i>Minimum pitch</i>	Main effect ***	Main effect ***	Main effect ***	Main effect ***
	H>A ***	H>A **	H>A **	H>A **
	H>N ***	H>N ***	H>N **	H>N ***
	A vs. N, <i>ns</i>	A vs. N, <i>ns</i>	A vs. N, <i>ns</i>	A vs. N, <i>ns</i>
<i>Maximum pitch</i>	Main effect ***	Main effect ***	Main effect ***	Main effect ***
	H>A **	H>A ***	H>A ***	H>A ***
	H>N ***	H>N ***	H>N ***	H>N ***
	A vs. N, <i>ns</i>	A>N **	A>N *	A vs. N, <i>ns</i>
<i>Pitch range</i>	Main effect **	Main effect ***	Main effect ***	Main effect ***
	H vs. A, <i>ns</i>	H>A ***	H>A ***	H>A ***
	H>N ***	H>N ***	H>N ***	H>N ***
	A vs. N, <i>ns</i>	A vs. N, <i>ns</i>	A>N *	A vs. N, <i>ns</i>
<i>Mean pitch</i>	Main effect ***	Main effect ***	Main effect ***	Main effect ***
	H>A **	H>A ***	H>A ***	H>A ***
	H>N ***	H>N ***	H>N ***	H>N ***
	A vs. N, <i>ns</i>	A>N **	A>N ***	A>N ***
<i>Pitch SD</i>	Main effect **	Main effect ***	Main effect ***	Main effect ***
	H vs. A, <i>ns</i>	H>A ***	H>A ***	H>A ***
	H>N **	H>N ***	H>N ***	H>N ***
	A vs. N, <i>ns</i>	A vs. N, <i>ns</i>	A vs. N, <i>ns</i>	A vs. N, <i>ns</i>
<i>Mean intensity</i>	Main effect **	Main effect ***	Main effect ***	Main effect ***
	H vs. A, <i>ns</i>	H>A ***	H>A ***	H>A ***
	H vs. N, <i>ns</i>	H>N **	H>N ***	H>N ***
	N>A **	A vs. N, <i>ns</i>	A vs. N, <i>ns</i>	A vs. N, <i>ns</i>

Notes: Line 1 of each cell refers to the main effect of emotion, and lines 2–4 show the results of pairwise comparisons. Significance levels of pairwise comparisons reflect a Bonferroni correction for multiple comparisons.

 $p < .001$;

**
 $p < .01$;

*
 $p < .05$.

Table 3

Mean (SD in parentheses) percentage of responses following each possible pattern for the first three gates;
H=Happy, A=Angry; N=Neutral

Response pattern	Percentage by participants	Percentage by items
NNN	36.1 (11.8)	36.2 (35.2)
NNA	2.3 (2.5)	2.2 (5.4)
NNH	1.0 (1.5)	1.0 (3.1)
NAA	5.9 (4.0)	6.0 (9.6)
NHH	3.8 (2.9)	3.7 (8.9)
NAN	0.4 (1.0)	0.4 (1.9)
NHN	0.5 (1.4)	0.5 (1.9)
NAH	0.1 (0.7)	0.1 (0.8)
NHA	0.1 (0.3)	0.0 (0.6)
HHH	16.9 (6.9)	16.8 (27.3)
HHA	0.6 (1.0)	0.6 (2.5)
HHN	0.6 (1.3)	0.6 (2.4)
HAA	1.7 (2.1)	1.7 (4.8)
HNN	1.1 (1.6)	1.1 (3.8)
HAH	0.1 (0.3)	0.1 (0.6)
HNH	0.1 (0.5)	0.1 (0.8)
HAN	0.0 (0.0)	0.1 (0.8)
HNA	0.0 (0.0)	0.0 (0.0)
AAA	22.7 (8.3)	22.8 (29.8)
AAN	1.8 (2.4)	1.8 (4.3)
AAH	0.4 (0.8)	0.4 (1.8)
ANN	2.2 (3.1)	2.2 (4.6)
AHH	0.9 (1.2)	0.9 (4.1)
AHA	0.2 (0.5)	0.1 (1.3)
ANA	0.3 (0.7)	0.3 (1.5)
AHN	0.1 (0.5)	0.1 (0.8)
ANH	0.2 (0.5)	0.2 (1.1)

Table 4

Mean number of times participants committed two types of errors (SD in parentheses)

<i>Incorrect label</i>			<i>Bias</i>	
<i>Happy</i>	<i>Angry</i>	<i>Neutral</i>	<i>Positive</i>	<i>Negative</i>
0.8 (1.1)	1.1 (1.6)	4.4 (2.6)	2.7 (2.1)	3.6 (2.6)
ns		$p = .000$	$p = .055$ (1-tailed)	
$p = .000$				

Table 5

Mean confidence ratings (1–5) and percent correct (SD in Parentheses) for happy, angry, and neutral sentences at Gates 1–5 and the final gate

	Happy	Angry	Neutral
Gate 1	2.0 (0.8)	1.9 (0.8)	2.1 (0.9)
	52% (18)	60% (19)	83% (13)
Gate 2	2.3 (0.9)	2.4 (0.9)	2.4 (0.9)
	62% (18)	74% (16)	87% (11)
Gate 3	2.6 (0.8)	2.7 (0.8)	2.8 (0.9)
	64% (19)	79% (13)	90% (10)
Gate 4	2.9 (0.8)	3.1 (0.8)	3.1 (0.8)
	66% (18)	81% (12)	92% (11)
Gate 5	3.1 (0.8)	3.3 (0.7)	3.4 (0.8)
	72% (16)	81% (12)	92% (11)
Final gate	4.2 (0.6)	4.2 (0.6)	4.4 (0.5)
	81% (13)	87% (11)	92% (10)

Table 6

Mean confidence ratings (1–5) for correct items only (SD in parentheses)

		Time ***	
		Isolation point	Final gate
<i>Emotion</i> [†]	Happy	2.07 (0.90)	4.17 (0.70)
	Angry	2.01 (0.90)	4.10 (0.70)
	Neutral	2.01 (1.09)	4.30 (0.64)

*Notes:**** Main effect of time, $p < .001$.[†] Main effect of emotion, *ns*.