Published in final edited form as:

Brain Cogn. 2009 February ; 69(1): 108–115. doi:10.1016/j.bandc.2008.06.002.

Lateralization of Visuospatial Attention Across Face Regions Varies With Emotional Prosody

Laura A. Thompson, Daniel M. Malloy, and Katya L. LeBlanc

Psychology Department, New Mexico State University

Abstract

It is well-established that linguistic processing is primarily a left-hemisphere activity, while emotional prosody processing is lateralized to the right hemisphere. Does attention, directed at different regions of the talker's face, reflect this pattern of lateralization? We investigated visuospatial attention across a talker's face with a dual-task paradigm, using dot detection and language comprehension measures. A static image of a talker was shown while participants listened to speeches spoken in two prosodic formats, emotional or neutral. A single dot was superimposed on the speaker's face in one of 4 facial regions on half of the trials. Dot detection effects depended on emotion condition—in the neutral condition, discriminability was greater for the right-, than for the left-, side of the face image, and at the mouth, compared to the eye region. The opposite effects occurred in the emotional prosody condition. The results support a model wherein visuospatial attention used during language comprehension is directed by the left hemisphere given neutral emotional prosody, and by the right hemisphere given primarily negative emotional prosodic cues.

Despite nearly 60 years of research on how human attention functions when processing electronic displays of text and graphic information or when understanding auditory speech, comparatively little is known about human attention as it applied to auditory-visual speech comprehension. The purpose of the present study is to investigate how a visuospatial attention mechanism that is used to process a talker's facial information is affected by the nature of prosodic information heard during a language comprehension task.

In face-to-face communication, information arrives to our language understanding systems from the talker's mouth, in the form of auditory linguistic and prosodic information, and from the face, in the form of visual articulatory cues (visible speech), facial expressions, eye gazes, and head movements. This smorgasbord of information available to our mind and senses is modulated by an attention system that can be flexibly adapted to suit the circumstances of the particular listening and viewing situation. Further, the amount of influence of visible speech is dependent on contextual and perceiver characteristics (e.g., Jordan & Sergeant, 2000; Sekiyama & Tohkura, 1993). Several studies conducted in our laboratory have yielded adult-age effects of visible speech influence during auditory-visual language processing. Compared to younger adults, older adults are usually found to be more reliant on visible speech (Thompson, 1995; Thompson & Malloy, 2004); although, the age effect reverses during extremely attention-demanding task situations, such as during a shadowing task, when younger

Address correspondence to: Laura A. Thompson, Department of Psychology, Box 30001/ MSC 3452/, New Mexico State University, Las Cruces, NM 88003, Phone: 505-646-4024, Fax: 505-646-6212, Email: E-mail: thompson@nmsu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

adults' shadowing performance is improved by the presence of visible speech while older adults' performance is not (Thompson & Guzman, 1999).

Two hypotheses for older adults' greater reliance on visible speech were tested (Thompson & Malloy, 2004). One hypothesis was that older adults have built up more finely tuned visible speech representations that are used during auditory-visible speech processing. This hypothesis was not supported. The hypothesis that was supported was that older adults focus more of their limited attention resources away from the eyes and toward the mouth area of the face during active speech processing. Greater focus on the mouth region compared to the eye region did not occur for either age group during silent video presentations of the same speech stimuli, suggesting that attention is not automatically directed to the mouth area as a consequence of movement in that spatial region. These findings suggest that the listener is choosing not to direct attention towards the mouth region in the absence of audible speech, even though the motion cues are present and available for processing in silent video conditions.

Attention has been long-argued to be both a data-driven process (a sequence of increasingly sophisticated analyses that begin with the incoming sensory data) and a conceptually-driven process (highest-level expectations of experiences further refined by analyses based upon a particular context) (Norman, 1976). One way to understand how the attention system functions during auditory-visual language understanding is to presume that it works simultaneously in a data-driven and a conceptually-driven manner. In a conceptually-driven manner, attention could be directed towards regions of the face that normally provide the maximal amount of sensory information to supplement the comprehension process. If attention is strategically directed towards regions of the talker's face using an attention mechanism that responds to the signal strength of the talker's facial display, attention should be greatest at the side of the face presenting the greatest amount of information.

There is evidence supporting this prediction. In human adults, greater movement (Wylie & Goodale, 1988) and amplitude of motion (Holowka & Pettito, 2002), and in nonhuman primates, earlier onset in the timing of movement (Hauser & Akre, 2001) appears on the left side of the face (viewer's right) during the production of emotional facial expressions. Importantly, the left side of photographic face images is perceived to be more expressive of facial emotion (Nicholls, Wolfgang, Clode, & Lindell, 2002). Research also shows a morphological asymmetry between the right and left sides of the mouth during the production of speech, such that the right side of the face, controlled by the adult talker's language (left) hemisphere, displays greater amplitude and velocity of motion than does the left side (Graves, Goodglass, & Landis, 1982; Wolf & Goodale, 1987). Human infants also show a right-side-wider effect when babbling (Holowka & Pettito, 2002). These researchers argue that this asymmetrical pattern of facial displays reflects an underlying asymmetry in activation of the two cerebral hemispheres, specifically, primary activation of the left hemisphere for linguistic productions and the right hemisphere for emotional expressions.

One study investigated visuospatial attention during auditory-visual language understanding using syllables as stimuli. Nicholls, Searle, and Bradshaw (2004) investigated asymmetries in the perception of auditory-visual speech with the McGurk effect (McGurk & MacDonald, 1976), an illusion which is produced when incongruent speech and visible speech are presented for syllable classification ("ga" mouth movements combined with an auditory "ba" results in the report of "da"). They covered up the right or left sides of the talker's face, and found that McGurk-type errors were lower when the side of the face with the most movement was covered (the right side), suggesting that the perceiver's attention is normally toward the left hemifield. Perhaps even more interesting is the fact that when the faces were mirror reversed, there were fewer McGurk-type errors when neither side of the face was covered. This suggests that participants were *directing* their attention to the side of the face that they expected would

convey the most information, i.e., the right side of the talker's face, which was actually the left side when mirror reversed. This finding implies that it is not the presence or absence of the facial cues that always elicits attention in a bottom-up fashion; instead, the *expectation* of the cues directs attention, at least in some circumstances, supporting a top-down or conceptually-driven model of attention.

While the Nicholls, et al.(2004) results support our assumption of the conceptually-driven nature of visuospatial attention during auditory-visual speech understanding (Thompson & Malloy, 2004), research conducted in our laboratory shows the opposite effect for the region of the face attended to—specifically, our research shows a visuospatial attention bias towards the talker's left side (right-side-of-face image) during visual-spoken language understanding (Thompson, Malmberg, Goodell, & Boring, 2004; Thompson & Malloy, 2004). Using a novel methodology, we previously investigated the distribution of visuospatial attention across a talker's face during the comprehension of lengthy speech passages (Thompson, et al. 2004). Dots were superimposed on the talker's face for 16 msec. Participants completed a comprehension task (the primary task), and the secondary task of dot detection. Dot detection performance for different areas of the face (right, left, above eyes, below mouth) indicated the degree of visuospatial attention on those areas. Our results showed greater attention was distributed on the right side of the talking face image, compared to the left side. Smeele, Massaro, Cohen, and Sittig (1998) also found a right visual field advantage in visible speech syllable and nonspeech (nonemotional) mouth movement identification, indicating primary involvement of the left hemisphere during the encoding of both types of mouth movements.

The discrepant right- versus left-side-of-face-image attention results could be partially explained by the nature of the tasks. Syllable classification of visual-spoken utterances containing often incongruent auditory and visual cues (Nicholls, et al. 2004) is quite different from our dual-task paradigm using lengthy speech passages of congruent auditory-visual stimuli. In syllable classification tasks, the perceiver may employ an atypical strategy of concentrating visuospatial attention on the most informative facial region for making a perceptual judgment, while thinking rather minimalist linguistic thoughts. In contrast, the dual-task comprehension paradigm is more similar to natural language understanding wherein all levels of linguistic and semantic analysis occur, albeit with the added secondary task of detecting a dot on someone's face. Thus, in our methodological paradigm the language hemisphere should be highly activated and the visuospatial attention system highly engaged. Regardless, results from our laboratory using lengthy speech stimuli in a comprehension task do not support the notion that a visuospatial attention mechanism is conceptually-driven to focus on the side of the face with the most informative visible speech cues, while the Nicholls et al. (2004) results using McGurk-type stimuli and a syllable identification task do support it.

In addition to the right-side-of-face-image effect, we found that greater attention was paid to the mouth than the eye area (Thompson, et al. 2004; Thompson & Malloy, 2004). We also showed a reversal of this effect, greater attention to the eye area than mouth area, based on differences in the emotional expressiveness of the talker (Thompson, et al. 2004). Perceivers might devote more attention to the eye area in attempt to discern emotion cues from the face (Ekman, 1997; Kimble, Forte, & Yoshikawa, 1981).

How can these differences in attention biases for different facial regions that occur during auditory-visual discourse processing be accounted for with a model of the distribution of attention during varying conditions of emotion interpretation?

There are many factors to take into consideration. To begin, it is customarily assumed that both hemispheres of the brain are actively engaged in processing information at all times. However, there exists a division of labor between the two coordinated brain hemispheres. One way in

which a division of labor functions is in dividing the visual input from the left and right visual fields for initial processing of the stimulus. In their investigation of visuospatial attention, Reuter-Lorenz, Kinsbourne, and Moscovitch (1990) used a line bisection task paradigm in which participants viewed a tachistoscopically-presented horizontal line and judged the location of a vertical intersect. Results showed attention was biased in the direction contralateral to the activated brain hemisphere, and referred to this general attention asymmetry as the attention-orienting hypothesis. The rightward- and leftward-attentional biases were also found not to depend on task relevance, suggesting a data-driven or automatic component to lateralized asymmetry of attention. Moreover, leftward and rightward attention asymmetries were found not to be a result of hemispatial position of the stimulus, in other words, orienting to an absolute region of space, but rather, results supported the notion of orienting in a direction of space. Thus far, the attention-orienting hypothesis has not been tested in the context of visuospatial attention during auditory-visual discourse processing. The results from Nicholls et al. (2004) in the non-mirror-reversal condition are consistent with the activation-orienting hypothesis, assuming greater dynamic motion produces a greater level of activation in the corresponding hemisphere, and, that attention is directed contralateral to the more highly activated hemisphere. However, the leftward bias produced even in the mirror-reversal condition highlights the fact that attention is modulated by a mechanism that is not purely data-driven.

Lateralization of emotion and language processing: Behavioral and brain evidence

Since the task in the present study incorporates a static face stimulus in conjunction with auditory speech, some consideration must be given to attention biases that might occur as a result of face processing. Many studies show a left visual field (right hemisphere) advantage for perceptual processing of static images of faces using tasks that do not involve language processing, but do involve emotion processing or making attractiveness judgments (e.g., Burt & Perrett, 1997; Klein, Moscovitch, & Vigna, 1976; Moscovitch, Scullion, & Christie, 1976). Topographic mapping of the brain shows that the right hemisphere activates when interpreting facial expressions of emotions (Gunning-Dixon, et al., 2003; Levy, Wagner, & Luh, 1990; Mikhailova, Davydov, & Morgunkova, 1996). Finally, Prodan, Orbelo, Testa, and Ross (2001) found that the right hemisphere was more highly activated when processing the upper region of the face while the left hemisphere was more highly activated while processing the lower region of the face. One explanation for this finding is that greater activation of the right hemisphere causes greater attention toward the pickup of emotion cues from the eye region. Likewise, differential activation of the left hemisphere could imply greater attention to the mouth for visible speech encoding.

Yet, since the task in the present study is primarily a language comprehension task, greater consideration should be given to patterns of attention that should be produced given varying types of auditory language input. Research going back to the 1960s using dichotic listening tasks has demonstrated a consistent right ear advantage (REA) in reporting words using both auditory-alone (e.g., Broadbent & Gregory, 1965; Hiscock, Inch, & Kinsbourne, 1999), and auditory-visual (Thompson, Garcia, & Malloy, 2007; Thompson & Guzman, 1999) speech stimuli, reflecting left hemisphere specialization for linguistic processing. Grimshaw, Kwasny, Covell, and Johnson (2003) recently reported evidence in a dichotic listening paradigm favoring a direct access interpretation; that is, each hemisphere was found capable of linguistic processing, however, the left hemisphere processed the auditory stimuli faster and more accurately. Topographic brain mapping studies have also shown primary activation in the left hemisphere during silent lipreading (Calvert, et al. 1997) and visible speech processing (Campbell, De Gelder, & De Haan, 1996).

When auditory speech understanding includes a greater emphasis on the interpretation of emotional prosody however, a different pattern of hemispheric activation emerges. While prosody can convey both linguistic and emotional information, prosody is referred to here as the intonational pattern of the voice that has emotive connotations for the listener. Dichotic listening tasks in normal participants almost universally show a left ear advantage (LEA) for comprehension of emotional prosody (e.g., Bryden & MacRae, 1989; Erhan, Borod, Tenke, & Bruder, 1998; Grimshaw, 1998; Grimshaw, et al. 2003; Stirling, Cavill, & Wilkinson, 2000). Brain imaging and event-related brain potential (ERP) studies find activation for emotional prosody primarily in right-hemisphere brain regions (Buchanan, Lutz, Mirasazade, Specht et al., 2000; George, Parekh, Rosinsky, Ketter, et al., 1996; Erhan et al., 1998; Shapiro & Danly, 1985; Shipley-Brown, Dingwall, & Berlin, 1988). Recent functional-anatomic findings from stroke patients support the claim that emotional prosody is a dominant and lateralized right-hemisphere function (Ross & Monnot, 2008). Finally, when the same spoken messages are presented in different emotional tones of voice, a REA is observed when participants attend to the linguistic content, but a LEA is observed when attending to emotional prosody (Ley & Bryden, 1982). These studies all support right hemisphere specialization for emotional prosody processing, and further, the Ley and Bryden (1982) study affirms that the processing of emotional prosody contains a conceptually-driven component.

A lateralization model of visuospatial attention during auditory-visual discourse processing

In summary, the evidence overwhelmingly supports a model of attention orienting during face-to-face discourse understanding situations that presumes that attention is, at least in part, a conceptually-driven process (Ley & Bryden, 1982; Nicholls, et al. 2004; Thompson & Malloy, 2004). Taking it one step further, a model of visuospatial attention during auditory-speech discourse processing predicts that attention is directed to different regions of the talker's face, depending on the extent to which the listener is actively engaged in emotion processing. Specifically, absent the goal of attending to the emotional content of the message, the left hemisphere should be more highly engaged for language processing than the right hemisphere. Consequently, owing to the combined findings that the left hemisphere is primarily responsible for linguistic processing (Grimshaw, et al. 2003) and spatial attention is oriented to the contralateral hemisphere (Reuter-Lorenz, et al. 1990), the right side of the face image should receive greater focus of attention than the left side. Furthermore, greater attention to the mouth than the eye region should also occur with speech stimuli spoken with neutral prosody (Prodan, et al. 2001) even when that region of the face is devoid of visible speech cues, due to the habit of looking at that region and expecting to find relevant information to encode. However, the relative degree of activation of the right hemisphere should increase in the presence of emotional content presented in the face and voice (e.g., Ross & Monnot, 2008). It follows that, in a condition with a high degree of emotional prosody, attention should be directed to the region of the face contralateral to the right hemisphere, the left side of the face image. Because the eye region is particularly rich with cues to the emotional content of the speaker's meaning (Ekman, 1997), and is associated with right hemisphere activation (Prodan, et al. 2001), attention should also be more directed to this region than to the mouth region of the face when speech contains a generous degree of emotional prosody.

A brief description of the methodological paradigm and hypotheses

In the present study, we explore the question of where is visuospatial attention directed on the speaker's face when the emotional prosody conveyed in the voice is either high or low. To this end, we conducted a dual-task experiment, where participants listened to speeches as their primary task, and responded to the presence or absence of a dot presented in one of four pre-cued locations on the talker's face. To rule out attention capture to eye or mouth regions because

of motion cues, a static image of the talker was presented. Unlike our earlier experiments, head and eye movements were controlled in this experiment through the use of a chin rest and 8 msec dot presentations, in order to more precisely determine the pattern of attention distribution across regions of the talker's face. A signal detection theory (SDT) paradigm was also used in assessing visuospatial attention to tease apart sensitivity and biasing effects (Green & Swets, 1966). Consistent with the requirements of SDT, attention was pre-cued to one of four regions on the face. This design parameter ensured that there was ambiguity only in the presence versus absence of the signal, and not also in terms of the location of the signal. The semantic and emotional content of speeches was controlled by presenting the same speeches in both emotion and neutral expression prosodic format to different participants. Two-thirds of the speeches were presented under noisy conditions with the intent of motivating the perceiver to devote a large amount of attention resources to the primary task of language comprehension. Comprehension was assessed with multiple choice questions presented as text and dispersed throughout eight segments within each speech.

Support for the lateralization of visuospatial attention model would be seen: (a) in greater dot detection discriminability (higher D' primes) for the right side of the face image than for the left side, and for the mouth region than the eye region, in the neutral prosodic expression condition, and (b) in greater dot detection discriminability for the left side of the face image than for the right side, and for the eye region than the mouth region, in the emotional prosodic expression condition.

Method

Participants

Participants were 14 adults (M age = 23.5 years, *range* 19–29; 6 male, 8 female), recruited from introductory psychology courses. All participants were right-handed, as determined using the Edinburgh Handedness Inventory.

Stimuli and apparatus

Stimuli included six videotaped speeches presented by one 22-year old female of Hispanic ethnicity. Speech content was edited from published versions to an approximate length of fifteen minutes. Topics included "College women and body image" (Dobkin & Sippy, 2002); "Take back the night" (Stutemeyer, 1998); "Women's rights are human rights" (H. Clinton, 1995); John F. Kennedy's commencement address at American University (1963); "HIV/AIDS and violence against women" (Piot, 1999); and Oprah Winfrey's commencement address at Wellesley College (1997). Each speech was divided into thirty-two 20–25 sec segments. Articulation rate averaged 145 words per minute.

Speeches were video recorded using a Sony CCD-TRV615 Digital HandyCam, and the resulting audio tracks were dubbed onto the final digitized recordings. The reader of the speeches was instructed to convey as little emotion as possible for the Neutral condition readings. For the Prosodic Expression condition, she was instructed to convey exaggerated emotional tones while speaking, conveying anger, sadness, fear, and happiness at appropriate moments given the content of the speech. A single static image of the actress with a neutral facial expression, taken from the video recording, was used for video presentation. A black dot, positioned on the nose in the middle of the face, served as a fixation point. The fixation point was also on the screen prior to the presentation of the static image. Speeches were graphically digitized into DVD-R format and professionally mixed with one of three levels of auditory distraction. The Multi-Talker Babble 20 (Auditec of St. Louis) was used for auditory noise. The Multi-Talker Babble 20 consists of 20 speakers reading different passages at the same time, and none of the individual voices are intelligible. In creating distraction levels used

in the experiments, we began with stereo recordings of the speeches at an average level of 73.5 dbA (No Distraction level). For the Low distraction condition, the multi-talker babble was recorded at the 70 db level and mixed with the speeches. The High distraction level was a 73.5 db recording of the speeches combined with the multi-talker babble recorded 10 decibels higher than the speech. Each of the six speeches was randomly assigned to a distraction condition.

Videos were modified by a certified multimedia producer using professional graphics editing equipment to include one small white dot (0.2 cm in diameter) that was superimposed on a single static image of a talker's face (see Figure 1). Dots were graphically positioned on the video frames so they appeared in a consistent location each time: (1) the Position 1 dot was aligned in front of the ear on the talker's left side (right visual field from the viewer's perspective; 3.89 cm right of center); (2) the Position 2 dot was centered on the talker's chin (5.64 cm below center); (3) the Position 3 dot was aligned in front of the ear at the talker's right side (left visual field, viewer's perspective; 3.87 cm left of center); and, (4) the Position 4 dot was centered on the talker's forehead just above the eyes (4.12 cm above center).

Between each segment, participants were asked whether they perceived a dot, and then were asked to choose from a multiple-choice set of answers, shown visually on a computer screen, whether they comprehended the semantic information presented. The multiple-choice format was employed to ensure that participants understood the speech and were not listening primarily for gist. Items referred to specific information mentioned once during the passage, for example:

Professor Woodrow Wilson said that every college graduate should be a man of his nation and a man of his __? (a) pride (b) time (c) honor (d) family.

Each question was pilot tested by five undergraduate students who did not listen to the speeches, so that the experiment had proper foils to reduce the probability of guessing the right answers without comprehending the speech. The visual stimuli were presented on a 15-in Sony Trinitron UltraScan P780 monitor controlled by a Dell Dimension 4400 computer. Audio stimuli were presented through small stereo speakers located on either side of the monitor.

Design and procedure

The design used was a 2 (Emotion: Neutral, Prosodic Expression) X 4 (Dot Position: Position 1, Position 2, Position 3, Position 4) within-subjects design. Participants saw one speech in each of three distraction levels within each of the two emotion conditions, for a total of six 32-trial speeches. Distraction level was not included in the analysis, however. Half of the trials contained a dot presentation. There were 48 trials for the dot detection task in the neutral condition, and 48 for the prosodic condition, 12 for each dot position. The ordering of emotion condition was counterbalanced across participants.

Overhead lights were turned off to prevent glare. Participants were placed in a chair one foot away from the monitor, and their head was positioned in a chin rest directly in the middle of the monitor. Instructions included the statements:

Your primary task is to pay as much attention to the speeches as you can while ignoring the background noise because you will be quizzed in detail on the specifics of the speech. Near the end of each segment you may or may not see a flash that looks like a dot on the person's face. The dots will appear on the face at four different locations for only a very short period of time. The secondary task is to correctly identify whether or not you saw a dot on the talker's face. This task is not as important as comprehending the speech.

Before the beginning of the first, ninth, seventeenth, and twenty-fifth segments participants saw a cue on the screen as to which of the four locations a dot would or would not appear; i.e.,

“If you see a dot, it will appear in Position 4.” Four of the eight trials in each position contained a dot stimulus, and this event ordering was determined randomly. A practice speech comprised of four segments with each level of distraction noise was presented for familiarizing the participant to the task.

Participants were also told that they must look at the computer screen at all times. The experimenter sat in an adjacent chair to ensure that all participants were following procedure. Comprehension and dot detection responses were recorded using a Dell Dimension 8200 notebook computer located on the table nearby. After the completion of the experiment, demographic questions were asked and participants were debriefed.

Results

Dot detection analyses

Hits and false alarms were converted to D prime (d') scores and were submitted to a repeated measures Analysis of Variance (ANOVA) using Emotion (2 levels) and Dot Position (4 levels) as within-subjects variables. In the initial ANOVA there was a significant Emotion X Dot Position interaction ($F(3, 39) = 9.20$, $MSE = 2.04$, $p < .0001$), consequently two ANOVAs were run on d' data in the Neutral and Prosodic Expression conditions separately to clarify the pattern of results in these two conditions. The ANOVA for the Neutral condition showed a significant main effect for dot position ($F(3, 39) = 5.69$, $MSE = 1.58$, $p = .002$). As in Thompson et al. (2004), paired samples t-tests were conducted to better understand the source of the main effect for dot position. These contrasts compared d' scores across the left versus right sides of the face image, and across mouth versus eye regions. Discriminability scores were significantly higher at Position 2 ($d' = 3.34$; mouth area) compared to Position 4 ($d' = 1.71$; just above the eye area), $t(13) = 3.65$, $p = .003$, $p_{rep} = .984$, $d = 1.23$. In addition, results showed a marginally significant difference between Position 1 ($d' = 3.18$; viewer's right visual field, talker's left side) and Position 3 ($d' = 2.11$; viewer's left visual field, talker's right side) dot detection scores, $t(13) = 1.91$, $p = .08$, $p_{rep} = .919$, $d = .81$.

The Prosodic Expression ANOVA also revealed a significant main effect for dot position ($F(3, 39) = 6.56$, $MSE = 1.70$, $p = .001$). Planned paired samples t-tests showed that discriminability scores were significantly higher in Position 3 ($d' = 3.59$) compared to Position 1 ($d' = 1.96$), $t(13) = 3.16$, $p = .008$, $p_{rep} = .979$, $d = 1.18$. Discriminability was also significantly higher at Position 4 ($d' = 2.94$) compared to Position 2 ($d' = 1.64$), $t(13) = 3.34$, $p = .005$, $p_{rep} = .924$, $d = .83$. The pattern of D prime scores in the Prosodic Expression was opposite to the pattern found in the Neutral Expression condition.

To directly compare discriminability data across Neutral and Prosodic Expression conditions, separate paired t-tests were conducted for data at each dot position. D prime scores were significantly higher in the Neutral condition than the Prosodic Expression condition at Position 1 ($t(13) = 2.24$, $p = .043$, $p_{rep} = .950$, $d = .95$) and Position 2 ($t(13) = 3.34$, $p = .005$, $p_{rep} = .989$, $d = 1.33$), and were significantly higher in the Prosodic Expression than the Neutral condition at Position 3 ($t(13) = 2.54$, $p = .025$, $p_{rep} = .964$, $d = 1.04$) and Position 4 ($t(13) = 2.68$, $p = .019$, $p_{rep} = .907$, $d = .764$). The different pattern of D prime and Beta scores for Neutral and Prosodic Expression conditions can be seen in Table 1.

Criterion analyses

Criterion scores (β) were examined using a repeated measures ANOVA for the two emotion conditions, separately, with dot position as a within-subject variable, to determine if the threshold at which participants set a mean criterion for answering “yes”, that a signal (dot) was presented varied across emotion conditions. In the Emotion condition, there was a marginally

significant main effect for dot position, $F(3,39) = 2.41$, $MSE = 0.91$, $p = .08$. Results revealed that criterion was set more conservatively at Position 4 ($\beta = 1.76$), compared to Position 2 ($\beta = 1.21$), $t(13) 2.21$, $p = .05$, however, criterion scores did not vary across Positions 3 and 1. In the Neutral condition, there was also a significant main effect for dot position, $F(3,39) = 2.88$, $MSE = 0.91$, $p = .048$. Criterion was set more conservatively at Position 2 ($\beta = 2.17$) compared to Position 4 ($\beta = 1.15$), $t(13) 3.17$, $p = .007$. Since criterion scores were higher at locations where d' scores were also higher, the pattern of results for the discriminability measure was not complicated by considerations of criterion scores.

Speech comprehension

Averages were calculated on the primary task multiple choice comprehension responses separately for each participant for each condition. A repeated measures ANOVA was completed, using Emotion condition and Dot Position as within-subjects variables. The results show that the primary task of comprehending speech was unaffected by the location on the face where the dot appeared prior to the comprehension question, and did not interact with emotional prosody (p 's $> .05$).

Discussion

Using a modification of the dot detection paradigm (Thompson, et al. 2004), we investigated differential patterns of visuospatial attention occurring during the comprehension of discourse presented with neutral and emotional prosodic expression. Our findings were: (a) for neutral prosodic expression conditions, visuospatial attention performance was better for the right side of the face image than for the left side, and for the mouth, compared to the eye area; (b) conversely, for emotional prosodic expression conditions, attention performance was better for the left side of the face image than for the right side, and at the eye, compared to the mouth area; and, (c) compared across prosodic expression conditions, dot detection performance was better at the mouth area and at the right side of the face image for the neutral condition, while performance was better at the eye region and the left side of the face image for the emotional prosody condition.

Our findings support a model of visuospatial attention that proposes that attention is oriented to the side of the face image that is contralateral to the primary activated hemisphere. The right hemisphere has been found to be the primary hemisphere active during emotion processing of faces (e.g., Gunning-Dixon, et al., 2003; Mikhailova, et al. 1996), and for emotional prosody processing (e.g., Bryden & MacRae, 1989; Erhan, et al. 1998; Grimshaw, 1998; Grimshaw, et al. 2003; Stirling, et al. 2000). We argue that, when emotional prosody is high, attention is biased to the region of the face that is contralateral to the right hemisphere, specifically, the left side of the face image. When emotional prosody is low, the left hemisphere for linguistic processing is more highly activated than is the right hemisphere, and attention is biased in the direction contralateral to the language hemisphere—to the right side of the face image.

The pattern of attention orienting found in this study does not support a data-driven model of human visuospatial attention during face-to-face discourse understanding. Importantly, since a static face image was used in the present study, differential attention to face regions was not due to variations across regions that are produced by a moving facial stimulus (Nicholls, et al. 2004). Moreover, the pattern of attention we found cannot be said to result from expectations formed by the perceiver's habit of looking at the more informative face region. If visuospatial attention were simply strategically oriented to the side of the face with the richest visual cues, then under emotional prosodic conditions we would expect to see greater attention at the right side of the face image (which corresponds to left side of the talker's face) because that is where the most informative emotional cues are present (Hauser & Akre, 2001; Holowka & Pettito, 2002; Wylie & Goodale, 1988). Under neutral prosodic conditions, we would expect to see

greater attention on the left side of the face image (which corresponds to the right side of the talker's face) because the right side of the face contains more visual cues for language comprehension (Graves, et al. 1982; Holowka & Pettito, 2002; Wolf & Goodale, 1987). Our results showed that, in both the neutral and the emotional prosodic conditions, attention was directed opposite to the most informative side of a talker's face. Taken together, the leftward and rightward biases found in the present study are consistent with the claim that spatial attention is oriented to the side of the stimulus contralateral to the more highly activated brain hemisphere (Reuter-Lorenz, et al. 1990).

In addition, consistent with Prodan, et al. (2001), our findings show that activation of the right hemisphere as a result of emotional prosodic discourse processing also directs attention to the eye region, most likely for the pickup of emotion cues (Ekman, 1997). In contrast, when the emotional content is neutral, the left hemisphere is the primary hemisphere activated for the pickup of visible speech cues at the mouth region of the face (Calvert, et al. 1997; Campbell, et al. 1996). Whether the eye and mouth region findings are indicative of conscious direction of attention to those regions, or a more passive "spillover" of attention due to hemispheric activation, cannot be determined from this study.

The rightward attention bias in the neutral prosodic expression condition obtained in this study using a static face stimulus, and in our previous studies using talking video images (Thompson, et al. 2004; Thompson & Malloy, 2004) oppose results recently reported by Nicholls, et al. (2004). They found that the McGurk effect was attenuated when the right side of the talker's mouth was covered, indicating a left hemifield attention bias during auditory-visual syllable classification. The Nicholls, et al. (2004) results support the argument that the perceiver's attention is drawn to the side of the talker's face that expresses more amplitude and velocity of motion (e.g., Wolf & Goodale, 1987). What kind of account can explain both sets of findings? We argue that perceivers' attention is directed differently, depending on the demands of the task situation. If the objective is to correctly classify a syllable in a tricky syllable classification task, when often the auditory and visual components do not match, perceivers strategically focus their attention on the most informative area of the face. If, on the other hand, the objective is to comprehend several sentences of discourse, perceivers also direct their attention, but in a different way. Attention is concentrated inwardly, on the multiple levels of meanings which become activated in cascading fashion all over the brain, originating in the language hemisphere, and "spilling out" into the right hemifield. We argue that, as long as the left hemisphere is more highly activated than the right, both a mouth-region focus and a rightward attention bias will naturally occur in auditory-visual language comprehension situations (Thompson, et al. 2004).

While some researchers argue for a primarily right-hemisphere emotion center, specifically the right parietotemporal region for the evaluation and interpretation of emotional information (e.g., Borod, et al. 1992; Canli, et al. 1998; Heller, 1990), research also suggests laterality differences that depend on the valence of emotional interpretation and experience (Heller, 1993). It is suggested that left hemisphere frontal regions are primarily activated during positive (cheerful) emotion experience, while right hemisphere frontal regions are activated during negative (depressed) experience (Davidson, 1992). Similarly, the right hemisphere could also be activated as a result of emotional arousal. Van Strien and Boon (1997) measured heart rate in their study of the effects of emotional music on visuospatial attention. They found that negatively-valenced music altered the visual field advantage in a lexical decision task, but had no effect on heart rate, leading them to conclude that arousal was not a factor in right-hemisphere influence on word processing.

There are several aspects of the design of the study that may limit the generality of conclusions that can be drawn regarding hemispheric activation during language processing. First, our study

did not control for the valence of emotional prosodic expression. However, we conducted a post-hoc analysis of the valence of emotion presentations in this condition, resulting in 88% classifications of negatively-valenced expressions. Since the majority of the prosodic variations we used would activate the right hemisphere regardless of which hypothesis was correct, this should not be an issue in interpreting the results. However, future research could explore whether attention is directed to the right side of the face during both negatively- and positively-valenced prosodic discourse processing. While other results suggest that right hemisphere activation is not due to arousal (Van Strien & Boon, 1997), nevertheless, more research needs to be conducted using emotional prosodic discourse (rather than music) and other physiological measures in order to rule out arousal as a factor in right hemisphere activation. Second, while it has long been assumed that only a minute fraction of right-handed individuals show a pattern of right hemisphere dominance for language, recent research finds that a surprising 7.5% of right-handed individuals are right-lateralized for language (Knecht, et al. 2000), and that the prevalence of left-lateralization activation patterns for language comprehension is equal across genders (Frost, et al. 1999). Further research should be conducted on larger sample populations than was employed in the present study, screened on the basis of hemispheric dominance for language instead of hand dominance, to investigate if the pattern of findings found here is true when language is lateralized to the right hemisphere.

In complex, naturalistic, multimodal language comprehension situations, the human attention system makes adjustments according to the emotional content of the talker's discourse. Our results demonstrate that these adjustments are best predicted, not on the basis of the talker's visual cues, but on the basis of the primary hemisphere activated in the perceiver.

Acknowledgments

The research was funded by a NIH NIGMS grant (#3S06 GM008136) and the NIH MBRS RISE grant program (GM61222).

References

- Borod JC, Andelman F, Obler LK, Tweedy JR, Welkowitz J. Right hemisphere specialization for the identification of emotional words and sentences: Evidence from stroke patients. *Neuropsychologia* 1992;30:827–844. [PubMed: 1407497]
- Broadbent DE, Gregory M. Accuracy of recognition for speech presentation to the right and left ears. *Quarterly Journal of Experimental Psychology* 1964;16:359–360.
- Buchanan TW, Lutz K, Mirzazade S, Specht K, Shah NJ, Zilles K, Jancke L. Recognition of emotional prosody and verbal components of spoken language: An fMRI study. *Cognitive Brain Research* 2000;9:227–238. [PubMed: 10808134]
- Burt DM, Perrett DI. Perceptual asymmetries in judgements of facial attractiveness, age, gender, speech and expression. *Neuropsychologia* 1997;35:685–693. [PubMed: 9153031]
- Bryden MP, MacRae L. Dichotic laterality effects obtained with emotional words. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology* 1989;1:171–176.
- Calvert GA, Bullmore ET, Brammer MJ, Campbell R, Williams SCR, McGuire PK, Woodruff PWR, Iversen SD, David AS. Activation of auditory cortex during silent lipreading. *Science* 1997;276:593–595. [PubMed: 9110978]
- Campbell R, De Gelder B, De Haan E. The lateralization of lip-reading: A second look. *Neuropsychologia* 1996;34:1235–1240. [PubMed: 8951835]
- Canli T, Desmond J, Zhao Z, Glover G, Gabrieli J. Hemispheric asymmetry for emotional stimuli detected with fMRI. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience* 1998;9(14):3233–3239.
- Clinton, H. Women's rights are human rights. Sep 5. 1995 Retrieved January 23, 2003, from <http://www.gos.sbc.edu/c/Clinton.html>

- Davidson RJ. Emotion and affective style: Hemispheric substrates. *Psychological Science* 1992;3:39–43.
- Dobkin, R.; Sippy, S. The college women's handbook. 2002. Retrieved January 23, 2003, from <http://www.feminist.com/market/wombus/colttext2.html>
- Ekman, P. What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). New York, NY: Oxford University Press; 1997.
- Erhan H, Borod JC, Tenke CE, Bruder GE. Identification of emotion in a dichotic listening task: Event-related brain potential and behavioral findings. *Brain and Cognition* 1998;37:286–307. [PubMed: 9665747]
- Frost JL, Binder JR, Springer JA, Hammeke TA, Bellgowan PSF, Rao SM, Cox RW. Language processing is strongly left lateralized in both sexes: Evidence from functional MRI. *Brain* 1999;122:199–208. [PubMed: 10071049]
- George MS, Parekh PI, Rosinky N, Ketter TA, Kimbell TA, Heilman KM, Herscovitch MD, Post RM. Understanding emotional prosody activates right hemisphere regions. *Archives of Neurology* 1996;53:665–670. [PubMed: 8929174]
- Graves R, Goodglass H, Landis T. Mouth asymmetry during spontaneous speech. *Neuropsychologia* 1982;20:371–381. [PubMed: 7133377]
- Green, DM.; Swets, JA. Signal detection theory and psychophysics. Huntington, NY: Krieger; 1966.
- Grimshaw GM. Integration and interference in the cerebral hemispheres: Relations with hemispheric specialization. *Brain and Cognition* 1998;36:108–127. [PubMed: 9520310]
- Grimshaw GM, Kwasny KM, Covell E, Johnson RA. The dynamic nature of language lateralization: Effects of lexical and prosodic factors. *Neuropsychologia* 2003;41:1008–1019. [PubMed: 12667536]
- Gunning-Dixon F, Gur R, Perkins A, Schroeder L, Turner T, Turetsky B, Chan R, Loughhead J, Alsop D, Maldjian J, Gur R. Aged-related differences in brain activation during emotional face processing. *Neurobiology of Aging* 2003;24:285–295. [PubMed: 12498962]
- Hauser MD, Akre K. Asymmetries in the timing of facial and vocal expressions by rhesus monkeys: Implications for hemispheric specialization. *Animal Behavior* 2001;61:391–400.
- Heller, W. The neuropsychology of emotion: Developmental patterns and implications for psychopathology. In: Stein, N.; Leventhal, BL.; Trabasso, T., editors. *Psychological and biological approaches to emotion*. Hillsdale, NJ: Erlbaum; 1990. p. 167–211.
- Heller W. Neuropsychological mechanisms of individual differences in emotion, personality, and arousal. *Neuropsychology* 1993;7:476–489.
- Hiscock M, Inch R, Kinsbourne M. Allocation of attention in dichotic listening: Differential effects on the detection and localization of signals. *Neuropsychologia* 1999;13:404–414. [PubMed: 10447301]
- Holowka S, Pettito LA. Left hemisphere cerebral specialization for babies while babbling. *Science* 2002;297:1515. [PubMed: 12202820]
- Jordan TR, Sergeant P. Effects of distance on visual and audiovisual speech recognition. *Language and Speech* 2000;43:107–124.
- Kennedy, JF. Commencement address at American University. Jun 10. 1963 Retrieved January 23, 2003, from <http://cs.umb.edu/jfklibrary/j061063.htm>
- Kimble CC, Forte RA, Yoshikawa JC. Nonverbal concomitants of enacted emotional intensity and positivity: Visual and vocal behavior. *Journal of Personality* 1981;49:271–278.
- Klein D, Moscovitch M, Vigna C. Attentional mechanisms and perceptual asymmetries in tachistoscopic recognition of words and faces. *Neuropsychologia* 1976;14:55–66. [PubMed: 1272513]
- Knecht S, Deppe M, Dräger B, Bobe L, Lohmann H, Ringelstein EB, Henningsen H. Language lateralization in healthy right-handers. *Brain* 2000;23:74–81. [PubMed: 10611122]
- Levy J, Wagner N, Luh K. The previous visual field: Effects of lateralization and response accuracy on current performance. *Neuropsychology* 1990;28:1239–1249.
- Ley RG, Bryden MP. A dissociation of right and left hemispheric effects for recognizing emotional tone and verbal content. *Brain and Cognition* 1982;1:3–9. [PubMed: 6927552]
- McGurk H, MacDonald J. Hearing lips and seeing voices. *Nature* 1976;264:746–748. [PubMed: 1012311]

- Mikhailova ES, Davydov DV, Morgunkova AN. Interhemispheric asymmetry of visual evoked potentials in recognition of emotional facial expressions. *Human Physiology* 1996;22:92–98.
- Moscovitch M, Scullion D, Christie D. Early vs. late stages of processing and their relation to functional hemispheric differences in face recognition. *Journal of Experimental Psychology: Human Perception and Performance* 1976;2:401–416. [PubMed: 993745]
- Nicholls MER, Searle DA, Bradshaw JL. Read my lips: Asymmetries in the visual expression and perception of speech revealed through the McGurk Effect. *Psychological Science* 2004;15:138–141. [PubMed: 14738522]
- Nicholls MER, Wolfgang BJ, Clode D, Lindell AK. The effect of left and right poses on the expression of facial emotion. *Neuropsychologia* 2002;40:1662–1665. [PubMed: 11992654]
- Norman, DA. *Memory and attention: An introduction to human information processing*. New York: John Wiley & Sons; 1976.
- Piot, S. HIV/AIDS and violence against women. Mar 3. 1999 Retrieved January 23, 2003, from <http://www.unaids.org/whatsnew/speeches/eng.html>
- Prodan C, Orbelo DM, Testa JA, Ross ED. Hemispheric differences in recognizing upper and lower facial displays of emotion. *Neuropsychiatry, Neuropsychology, & Behavioral Neurology* 2001;14:206–212.
- Reuter-Lorenz PA, Kinsbourne M, Moscovitch M. Hemispheric control of spatial attention. *Brain and Cognition* 1990;12:240–266. [PubMed: 2340154]
- Ross ED, Monnot M. Neurology of affective prosody and its functional-anatomic organization in the right hemisphere. *Brain and Language* 2008;104:51–74. [PubMed: 17537499]
- Sekiyama K, Tohkura Y. Inter-language differences in the influence of visual cues in speech perception. *Journal of Phonetics* 1993;21:427–444.
- Shapiro BE, Danly M. The role of the right hemisphere in the control of speech prosody in prepositional and affective contexts. *Brain and Language* 1985;25:19–36. [PubMed: 4027566]
- Shipley-Brown F, Dingwall WO, Berlin CI. Hemispheric processing of affective and linguistic intonation contours in normal subjects. *Brain and Language* 1988;33:16–26. [PubMed: 3342317]
- Smeele PMT, Massaro DW, Cohen MM, Sittig AC. Laterality in visual speech perception. *Journal of Experimental Psychology: Human Perception Performance* 1998;24:1232–1242.
- Stirling J, Cavill J, Wilkinson A. Dichotically presented emotionally intoned words produce laterality differences as a function of localization task. *Laterality* 2000;5:363–371. [PubMed: 15513153]
- Stuteman, S. Take back the night address. Apr 22. 1998 Retrieved January 23, 2003, from <http://www.feminist.com/resources/artspeech/violence/stut.htm>
- Thompson LA. Encoding and memory for visible speech and gestures: A comparison between young and older adults. *Psychology and Aging* 1995;10:215–228. [PubMed: 7662181]
- Thompson LA, Garcia E, Malloy D. Reliance on visible speech cues during multimodal language processing: Individual and age differences. *Experimental Aging Research* 2007;33:373–397. [PubMed: 17886014]
- Thompson LA, Guzman FA. Some limits on encoding visible speech and gestures using a dichotic shadowing task. *Journal of Gerontology: Psychological Sciences* 1999;54B:347–349.
- Thompson LA, Malloy DM. Attention resources and visible speech encoding in older and younger adults. *Experimental Aging Research* 2004;30:1–12. [PubMed: 14660330]
- Thompson L, Malmberg J, Goodell N, Boring R. The distribution of attention across a talker's face. *Discourse Processes* 2004;38:145–168.
- Van Strien JW, Boon CA. Altered visual field asymmetry for lexical decision as a result of concurrent presentation of music fragments of different emotional valences. *Journal of the International Neuropsychological Society* 1997;3:473–479. [PubMed: 9322407]
- Van Strien JW, Heijt R. Altered visual field asymmetries for letter naming and letter matching as a result of concurrent presentation of threatening and nonthreatening words. *Brain and Cognition* 1995;29:187–203. [PubMed: 8573332]
- Winfrey, O. Oprah Winfrey's commencement address. May 30. 1997 Retrieved January 23, 2003, from <http://www.wellesley.edu/publicaffairs/pahomepage/winfrey.html>

- Wolf ME, Goodale MA. Oral asymmetries during verbal and non-verbal movements of the mouth. *Neuropsychologia* 1987;25:375–396. [PubMed: 3601043]
- Wylie DR, Goodale MA. Left-sided oral asymmetries in spontaneous but not posed smiles. *Neuropsychologia* 1988;26:823–832. [PubMed: 3194048]

Figure 1.

Instruction screen shown prior to a set of eight trials. Instruction above the face reads, “If you see a dot it will appear in Position 4” (or 1, 2, 3).

Table 1
Mean (SD) of D Primes, Betas, Hit Rate, False Alarm, and Comprehension Scores
At Each Dot Position For Neutral and Prosodic Emotion Expression Conditions

	Dot Position			
	1	2	3	4
Neutral Condition				
D Prime	3.18 (1.3)	3.34 (1.3)	2.11 (1.4)	1.71 (1.4)
Beta	1.76 (0.9)	2.17 (1.0)	1.51 (1.3)	1.15 (0.8)
Hit Rate	0.89 (0.1)	0.84 (0.1)	0.67 (0.2)	0.67 (0.2)
False Alarm Rate	0.08 (0.1)	0.07 (0.1)	0.19 (0.2)	0.19 (0.2)
Comprehension	82.2 (20.2)	80.9 (20.1)	83.5 (19.4)	85.0 (19.1)
Prosodic Expression Condition				
D Prime	1.96 (1.3)	1.64 (1.3)	3.59 (1.4)	2.94 (1.8)
Beta	1.42 (1.2)	1.21 (1.1)	2.11 (0.9)	1.76 (1.1)
Hit Rate	0.70 (0.1)	0.66 (0.1)	0.87 (0.1)	0.79 (0.2)
False Alarm Rate	0.18 (0.2)	0.21 (0.2)	0.07 (0.1)	0.12 (0.1)
Comprehension	84.5 (16.5)	83.8 (18.8)	83.5 (19.2)	82.4 (19.6)