Left Auditory Cortex and Amygdala, but Right Insula Dominance for Human Laughing and Crying

Kerstin Sander and Henning Scheich

Abstract

Evidence suggests that in animals their own species-specific communication sounds are processed predominantly in the left hemisphere. In contrast, processing linguistic aspects of human speech involves the left hemisphere, whereas processing some prosodic aspects of speech as well as other not yet well-defined attributes of human voices predominantly involves the right hemisphere. This leaves open the question of hemispheric processing of universal (species-specific) human vocalizations that are more directly comparable to animal vocalizations. The present functional magnetic resonance imaging study addresses this question. Twenty subjects listened to human laughing and crying presented either in an original or time-reversed version while performing a pitch-shift detection task to control attention. Time-reversed presentation of these sounds is a suitable auditory control because it does not change the overall spectral content. The auditory cortex, amygdala, and insula in the left hemisphere were more strongly activated by original than by time-reversed laughing and crying. Thus, similar to speech, these nonspeech vocalizations involve predominantly left-hemisphere auditory processing. Functional data suggest that this lateralization effect is more likely based on acoustical similarities between speech and laughing or crying than on similarities with respect to communicative functions. Both the original and time-reversed laughing and crying activated more strongly the right insula, which may be compatible with its assumed function in emotional self-awareness.

INTRODUCTION

Auditory processing of own species-specific vocalizations in mammals seems to engage predominantly the left hemisphere of the brain. Behaviorally, this was demonstrated by the preference of the right ear (i.e., the left hemisphere) for conspecific vocalizations, for example, in monkeys. Neither listening to meaningful bird vocalizations (Hauser & Andersson, 1994) nor to vocalizations of a different monkey species (Petersen et al., 1978) led to such a right ear advantage, although acoustic parameters of the vocalizations tested were equally well discriminated by different species of macaque monkeys (Petersen et al., 1984). Electrophysiological data support these findings. In the common marmoset monkey, neurons of the left primary auditory cortex (AC) responded selectively and with the largest discharge rate to natural species-specific vocalizations as compared to their time-reversed versions (Wang & Kadia, 2001; Wang, Merzenich, Beitel, & Schreiner, 1995). These original and time-reversed monkey vocalizations did not exhibit such neuronal response differences in the left primary AC of cats, which shares basic physiological properties with that of the marmoset monkey (Wang & Kadia, 2001). In the squirrel monkey, extracellular recordings of neurons in the primary and secondary AC did not reveal hemisphere differences in response to original or time-reversed versions of conspecific vocalizations (Glass & Wollberg, 1983). Nevertheless, neural activity in the left AC was stronger in primary than in secondary areas, but not in the right AC. Interestingly, the house mouse also preferred the right ear for processing species-specific communication sounds like its pups’ ultrasounds (Ehret, 1987). C-Fos immunocytochemistry revealed that the left dorsoposterior AC of female house mice exhibited the largest number of Fos-positive cells in response to natural mouse pup calls as compared to the right hemisphere and to artificial versions of the pup calls (Geissler & Ehret, 2004). Thus, most available results in animals suggest rather a left-hemisphere involvement of processing species-specific vocalizations, and some discrepancies of results might reflect methodological differences of sampling or of direct and indirect measures of neural activity related to auditory processing.

In humans, differences between the left and right hemisphere are observed in processing speech and nonverbal human vocalizations. Similarly, differences are found for processing linguistic (conceptual) and nonlinguistic aspects of speech (affective prosody, voice quality in general). Dichotic listening tasks revealed a right ear (left hemisphere) advantage for spoken words (Kreiman & Van Lancker, 1988; Kimura, 1964) and the
opposite, a left ear (right hemisphere) advantage for melodies, including hummed melodies, and for affective nonverbal vocalizations like human laughing and crying (King & Kimura, 1972; Kimura, 1964). Supporting evidence of this hemispheric division of labor in humans is provided by recent brain imaging studies using functional magnetic resonance imaging (fMRI) or positron emission tomography. Left lateralization of neural activity was revealed in response to speech contrasted with either its “frequency mirrored” (Narain et al., 2003) or backward versions (Crinion, Lambon-Ralph, Warburton, Howard, & Wise, 2003; but see Kimura & Folb, 1968) or with resting conditions (Binder et al., 2000). In contrast, anterior parts of the right superior temporal gyrus near the superior temporal sulcus were selectively engaged by comparing a choice of human voices with a single voice (Belin & Zatorre, 2003). Right lateralization of neural activity was also elicited by emotion identification in sentences and by expressiveness judgments of sentences spoken with different affective prosodies (Wildgruber, Pihan, Ackermann, Erb, & Grodd, 2002).

As lesion studies demonstrate, perception, comprehension, and production of affective prosodies seem to be heavily based on an intact right brain hemisphere (for reviews, see Myers, 1999; Joannette, Goulet, & Hannequin, 1990). However, it does not seem possible to predict any specific prosodic impairment with respect to lesion site and extent in the right hemisphere, as cortical and subcortical as well as anterior and posterior lesions may lead to prosodic impairments.

From this preliminary evidence, it would appear that in animals species-specific communication sounds are processed predominantly in the left hemisphere. In contrast, processing linguistic aspects of human speech involves the left hemisphere, whereas processing some prosodic aspects of speech as well as other not yet well-defined attributes of human voices predominantly involve the right hemisphere. This leaves open the question of hemispheric processing of transculturally universal (species-specific) human vocalizations that are more directly comparable to animal vocalizations. Such vocalizations with strong emotional valence are, for example, laughing and crying.

To adjust the design of the investigation to the cited monkey studies and to some imaging studies in humans, we analyzed activations by listening to laughing and crying in comparison to the backward versions of the same stimuli. From the point of view of stimulus properties, the backward presentations served as a certain type of acoustical control because the only difference between original and time-reversed versions was their short-term spectrotemporal dynamics, whereas their overall spectral contents and a sequential structure of sound elements were preserved. The backward versions still sounded laughing- or crying-like but distinctly strange.

Using fMRI, we have recently demonstrated that listening to laughing and crying causes strong bilateral activation of the human amygdala, insula, and AC (Sander & Scheich, 2001). These brain structures are not only anatomically interconnected (Yukie, 2002; Mesulam & Mufson, 1985), but all participate in the processing of emotional acoustic stimuli (Bamiou, Musiek, & Luxon, 2003; Phillips et al., 1998; Habib et al., 1995). Therefore, the present study focused on the amygdala, insula, and AC. Similarly to Kling, Lloyd, and Perryman (1987), who demonstrated that evoked field potentials in the monkey amygdala exhibited stronger power while stimulated with species-specific vocalizations than with control stimuli (pure tone of 400 Hz), we expected stronger amygdala activation in response to laughing and crying than to their time-reversed versions.

Twenty right-handers were scanned in an MR scanner while listening to laughing and crying in an original and time-reversed manner, that is, forward and backward. As a nonspecific attention control task, they had to detect randomly introduced upward shifts of spectral content (pitch shifts) in the stimuli. Subjects were debriefed on emotional stimulus impacts after scanning and had to report on perceived naturalness and nonnaturalness of stimuli.

Acoustically responsive structures like AC may contain feature maps in which stimuli are spatially represented and different stimuli lead to different spatial representations (Schulze, Hess, Ohl, & Scheich, 2002). Blood oxygen level dependent (BOLD) signal intensity alone as a measure to compare responsiveness to different sounds would only be sufficient if spatial representations are the same. Therefore, signal intensity is complemented in the present study by the measure of activated voxels and intensity-weighted volumes (IWVs). For the same reasons of unclear spatial overlaps of representations, BOLD responses to a given sound were compared to a resting condition and BOLD responses to different sounds were directly compared (direct contrast) (see Discussion).

RESULTS

Behavioral Data

The mean detection rate of correctly identified pitch shifts across all stimuli was 78.7% (SEM = 5.8). Detection rate was significantly better during forward than backward stimulus presentations (mean ± SEM: forward = 82.5 ± 5.6%, backward = 74.9 ± 6.5%, p = .048). Detection rates did not differ between “laughing” and “crying” or between “backward laughing” and “backward crying” (both comparisons, p > .3).

Possible influences of subjective mood changes on emotional stimulus perception were controlled via questionnaires directly filled out before and after the experiment. Apparently, there was no change in the subjects’
mood upon perception of laughing and crying either in forward or backward presentation mode. Scores of the mood scales "good mood/bad mood" and "rest/restlessness" did not differ before and after the experiment (both comparisons, $p > .2$), whereas subjects felt somewhat tired afterward ($p = .027$). However, the mean score of the scale "awakeness/tiredness" (12.5) was still slightly above the scale's mean (12).

Subjective ratings of the pleasantness, emotionality, intensity, and naturalness of the stimulus were analyzed for those 11 subjects who immediately reported having perceived a difference between forward and backward presentations of laughing and crying (a question as to whether the presented stimuli were in any way different from each other) (Figure 1). This does not mean, however, that the remaining nine subjects did not hear a difference between forward and backward stimulus presentations. Most of these subjects, when recapitulating the stimuli, realized a difference between them, but reported that during the experiment they were concentrating on the task and not much on possible stimulus differences.

Laughing was rated the most pleasant and most intense stimulus, differing significantly from crying, backward laughing, and backward crying (pleasantness comparisons, $p \leq .018$; intensity comparisons, $p \leq .025$). Laughing was rated similarly to crying with respect to emotionality and naturalness (both comparisons, $p > .2$), but was perceived as more emotional and more natural than the two backward stimulus presentations (emotionality comparisons, $p \leq .031$; naturalness comparisons, $p \leq .004$). Interestingly, crying was rated significantly more pleasant and more emotional than backward crying (both comparisons, $p \leq .016$) but not backward laughing (both comparisons, $p > .1$). Crying was rated more natural than backward crying and backward laughing (both comparisons, $p = .002$) but did not differ in intensity rating from the backward presentations (both comparisons, $p > .29$). Ratings of the four stimulus qualities did not differ between backward laughing and backward crying (all comparisons, $p > .18$). To summarize, laughing and crying led to different pleasantness and intensity ratings in accordance with their emotional valence and acoustical characteristics (intensity). However, backward playing of laughing and crying eliminated rating differences between them. They were now perceived as equally unpleasant, not emotional, being of middle intensity, and clearly unnatural.

**fMRI Data**

The analysis first covered all activation differences obtained by the four stimuli. Further analysis focused on differences between forward and backward presentations of stimuli and thus corresponded to the subjective classification of naturalness and nonnaturalness. For this purpose, the difference in emotional valence between laughing and crying was ignored. Third, neural activity was compared between those subjects who spontaneously reported having heard differences between forward and backward stimulus presentations (D group) and those who did not (ND group). This comparison attempted to reveal whether a spontaneous recognition of differences in stimulus presentation mode alters neural activity.

Functional activation was analyzed by correlation analysis and carried out with the custom-made software package KHORFu (Gaschler, Schindler, & Scheich, 1996), whereas potential differences between hemispheres, stimuli, and groups were tested by using analyses of variance with repeated measurements (see Methods).
Amygdala

Laughing and crying, presented forward or backward, elicited bilateral amygdala activation compared to resting condition (see Methods), but not to the same degree. Stronger left than right amygdala activation was elicited by crying (Figure 2A). In contrast, listening to backward laughing produced the reverse, that is, stronger right than left amygdala activation. Hemispheric differences for laughing and backward crying were similar to those of their counterparts although to a lesser degree. In the left amygdala, laughing and crying caused stronger activation than their backward versions. In the right amygdala, laughing and backward laughing elicited a similar amount of activation, whereas crying led to slightly less activation than backward crying. Thus, only in the left amygdala were differences of activation found, which point to a differentiation of natural and unnatural stimuli. The effects of the factors hemisphere and stimulus were not significant ($p > .4$).

The analysis of stimulus naturalness without differentiating between laughing and crying showed stronger left amygdala activation for the forward versions (IWV and number of voxels, $T = 2.192, df 19, p = .041$) which also caused stronger left than right activation (IWV and signal intensity, $T = 2.149, df 19, p = .045$) as reflected by the interaction of Hemisphere × Stimulus (all measures, $F = 6.674, df 1/18, p = .029$) (Figure 2B). The stronger right than left amygdala activation in response to backward stimulus presentations did not reach significance ($p > .05$).

Across both analyses, the stronger activation of the left amygdala by natural forward versions of stimuli was independent of perceived or nonperceived naturalness of stimuli (D and ND groups). However, activation was stronger in the nonperceived (ND) group (all measures, $F = 5.593, df 1/18, p = .019$).

The amygdala was the only structure analyzed in which signal intensity for natural stimuli was stronger in the left than in the right hemisphere (see above). Activation differences by direct stimulus contrasts, laughing versus backward laughing and crying versus backward crying, were not found.

Insula

Laughing and crying, presented either forward or backward, led to stronger right than left insula activation above resting condition (IWV and number of voxels, $F = 6.435, df 1/18, p = .001$) (Figure 3A).

Figure 2. Mean amygdala activation (number of voxels) ± SEM above resting condition. (A) Activations in response to laughing and crying, presented forward or backward, and (B) collapsed activations by laughing and crying (Forward) and their backward versions (Backward) are shown in the left and right hemisphere.

Figure 3. Mean insula activation (number of voxels) ± SEM above resting condition. (A) Activations in response to laughing and crying, presented forward or backward, and (B) collapsed activations by laughing and crying (Forward) and their backward versions (Backward) are shown in the left and right hemisphere.
Collapsing activation by laughing and crying showed stronger right than left insula activation (IWV and number of voxels, $F = 13.485$, $df = 1/18$, $p = .002$) (Figure 3B). The interaction of Hemisphere $\times$ Stimulus $\times$ Group was significant (IWV and number of voxels, $F = 8.403$, $df = 1/18$, $p = .010$) and revealed that the right insula activation dominance was significant for backward (IWV and number of voxels, $T = −3.455$, $df = 19$, $p = .003$) and not for forward stimulus presentations ($p = .7$). Like the amygdala, the left insula was more strongly activated by forward than backward stimulus presentations (number of voxels $T = 2.090$, $df = 19$, $p = .050$). In contrast, the right insula was more strongly activated by backward than forward stimulus presentations (IWV, $T = −2.294$, $df = 19$, $p = .033$).

The global pattern of insula activation was not influenced by subjects’ reports on naturalness of stimuli. Similar to the amygdala, overall insula activation was stronger in the ND than the D group (both analyses: IWV, $F = 7.496$, $df = 1/18$, $p = .014$). Collapsing activation by laughing and crying showed that this group effect was significant in the right insula for backward (IWV and number of voxels, $T \leq −2.703$, $df = 18$, $p = .015$) but not for forward stimulus presentations.

The direct stimulus contrasts produced only a few significant voxels without any hemispheric differences.

**Auditory Cortex**

Laughing and crying, forward and backward, elicited bilateral activation of the AC above resting condition (Figure 4), globally (across auditory territories) and in each of the four different functional–anatomical territories, TA, T1, T2, and T3 (Figure 6B, see Methods). In TA (IWV and signal intensity, $F = 3.170$, $df = 3/54$, $p = .032$) and across auditory territories (signal intensity, $F = 3.668$, $df = 3/54$, $p = .018$), the main effect of stimulus was significant, but none of the post hoc comparisons. In TA, the interaction of Stimulus $\times$

---

**Figure 4.** Stimulus versus resting condition contrasts. Mean AC activation (number of voxels) ± SEM above resting condition. (A) Activations in response to laughing and crying, presented forward or backward, and (B) collapsed activations by laughing and crying (Forward) and their backward versions (Backward) are shown for each territory (TA, T1, T2, and T3) and globally, across the four territories (total_left, total_right), in the left and right hemisphere. For definition of AC territories, see Figure 6B.
Group was significant (IWV, $F = 3.088$, $df\,3/54$, $p = .035$), without any significant post hoc comparison. No effect of the factor hemisphere was observed ($p > .8$).

Collapsing activation by laughing and crying (Figure 4B) revealed in T1 and T2 significant interactions of Hemisphere $\times$ Stimulus (signal intensity, $F = 4.439$, $df\,1/18$, $p = .049$) and Hemisphere $\times$ Stimulus $\times$ Group (signal intensity, $F = 5.075$, $df\,1/18$, $p = .037$), respectively, without any significant post hoc comparison. In T3, the interaction of Stimulus $\times$ Group was significant (IWV, $F = 5.126$, $df\,1/18$, $p = .036$), but none of the post hoc comparisons.

However, directly contrasting AC activation by laughing versus backward laughing and crying versus backward crying revealed a strong left lateralization (Figures 5A and 6C and D). This was significant in T2 (IWV and number of voxels, $F = 5.451$, $df\,1/18$, $p = .031$) and across all territories of AC (number of voxels, $F = 6.157$, $df\,1/18$, $p = .023$) (Figure 5A). In T2, the interaction of Stimulus $\times$ Group was significant (IWV, $F = 7.041$, $df\,1/18$, $p = .016$), but none of the post hoc comparisons.

In either hemisphere, the laughing and crying forward versus backward contrasts did not differ. Collapsing activation contrasts by laughing and crying revealed significant left lateralization in T2 (IWV and number of voxels, $F = 4.486$, $df\,1/18$, $p = .048$) and in T3 as well as across all territories (number of voxels, $F = 5.470$, $df\,1/18$, $p = .031$) (Figure 5B).

Naturalness recognition of stimuli did not alter the global AC activation pattern for each of the four stimuli above resting condition. Direct stimulus contrasts revealed left lateralization of activity for both the D and ND groups. Contrary to the amygdala and insula, activation in the AC was stronger for the D than the ND group (number of voxels in TA, T3, and across territories, Figure 5).

**Figure 5.** Direct stimulus contrasts. Mean AC activation (number of voxels) $\pm$ SEM in response to directly contrasting forward versus backward presentations of laughing and crying. (A) Activations by the contrasts laughing versus backward laughing (L_vs_BL) and crying versus backward crying (C_vs_BC), and (B) by collapsing activation contrasts by laughing and crying shown for each territory (TA, T1, T2, and T3) and globally, across the four territories (total_left, total_right), in the left and right hemisphere. For definition of AC territories, see Figure 6B.
DISCUSSION

Influence of Time-reversed Laughing and Crying on Behavior

Listening to laughing and crying either in original or in time-reversed version had a clear-cut effect on behavior. The detection rate of upward shifts in spectral content was better during laughing and crying than during their backward versions. This is in accordance with identification results obtained for spoken stories and for instrument sounds played forward or backward (Crinion et al., 2003; Paquette & Peretz, 1997).

Subjective stimulus ratings showed that laughing and crying were perceived as intended. In accordance with their emotional valence and acoustic characteristics, laughing was rated as more pleasant and more intense than crying, whereas both were rated as similarly emotional and natural. Backward laughing...
and backward crying were both perceived as equally unpleasant, not emotional, and unnatural. Although laughing still sounds laughing-like when played backward, acoustical analyses of adult laughing showed that the decrease in amplitude of single laugh segments across laughter, a decrescendo, is reversed to a crescendo by presenting laughing backward (Provine & Yong, 1991). Furthermore, the higher density of laugh segments at the beginning of laughter as well as their fundamental frequency contours are influenced by presenting laughing in a time-reversed manner (Bachorowski, Smoski, & Owren, 2001). To illustrate the segmental structure as well as the formant-like frequency contours, one of the female laughter segments used in the present study is depicted in Figure 6E (Boersma & Weenink, 1992–2003). To our knowledge, there are no comparable data available on similar acoustical analyses of adult crying. However, acoustical analyses of different classes of healthy infant cries revealed that pain and hunger cries are both characterized by a specific asymmetrical time course of spectral change (Wasz-Höckert, Lind, Vuorenkoski, Partanen, & Valanne, 1968). This means that representing such cries backward sounds unfamiliar or even strange to listeners. Thus, although the overall spectral contents of laughing and crying and their backward versions were identical, changes of spectro-temporal dynamics on a fine and on a coarse scale by time-reversed sound presentations may have caused the decline in pitch-shift detection performance and the change in subjective stimulus ratings, presumably reflecting the subjects’ acoustical nonfamiliarity with such time-reversed sounds.

**Influence of Time-reversed Laughing and Crying on Neural Activity**

**Stimulus versus Resting Condition Contrasts**

Listening to laughing and crying in both original and time-reversed versions led to bilateral activations of the amygdala, insula, and AC when neural activity was determined by contrasting stimuli with resting condition. With respect to hemispheric differences, stronger left than right amygdala activation was observed in response to original laughing and crying, and the reverse, stronger right than left amygdala activation in response to time-reversed stimuli. In the insula, both original and time-reversed stimuli elicited stronger right- than left-hemispheric activation, whereas in the AC, no activation lateralization was observed. The left amygdala differentiated between original and time-reversed laughing and crying with stronger activation by the original stimuli. Similar activation dominance by natural stimuli was observed in the left insula, whereas the reverse pattern, activation dominance by time-reversed laughing and crying, was observed in the right insula. No stimulus differentiation was observed in AC activity when referred to resting condition.

**Direct Stimulus Contrasts**

When activity differences were determined by directly contrasting natural with time-reversed laughing and crying, no activation at all was detected in the amygdala, and only a few activated voxels were detected in the insula without any hemisphere or stimulus differences. In the AC, a strong left lateralization of activation was observed, which was most prominent in the nonprimary posterior AC territories T2 (centered at first Heschl’s sulcus and extending in anterior parts of the planum temporale) and T3 (posterior parts of the planum temporale). It was shown for contrasting activations by laughing versus backward laughing and crying versus backward crying, but was stronger for the laughing contrast as well as for the collapsed activation contrast of natural versus time-reversed presentations of both stimuli together, that is, irrespective of emotional valence.

**Implications of Different Activation Contrasts**

BOLD activation determined by contrasting stimulus conditions and resting condition and by contrasting different stimulus conditions directly may theoretically lead to completely different results. This will depend on whether stimuli generate spatial representations in neural maps; that is, sound differences may generate different distinct spatial representations even though the activities for the two representations may be similar. In this case, direct stimulus contrast would reveal a difference, but not stimulus activations referred to resting condition. This seems to be the case in the AC. By contrast, in structures where stimulus differences are not spatially mapped or mapped only weakly, leading to largely overlapping representations, direct stimulus contrasts may not show differences. However, if two spatial representations have a large intersection set of voxels and one is slightly larger than the other, contrasts of stimuli versus resting condition may show differences. This may be the case in the amygdala and insula.

A difference between forward and backward is identified by the cortex as shown by direct contrast criterion and that difference is identified in the left AC and does not depend on reports of heard differences between stimuli. Independent of consciously recognizing stimulus differences, the left AC seems to “realize” naturalness.

Activation contrasts of the left AC (laughing vs. backward laughing, crying vs. backward crying) were not caused by so-called deactivations as another possibility of generating direct activation contrasts, that is, when one stimulus generates activation and another stimulus of similar spatial representation generates deactivation (Gusnard & Raichle, 2001). No negative correlations...
between the trapezoid correlation vectors and the time series of mean gray value changes of voxels were observed in the AC. Therefore, it is more likely that in the various maps of AC original and time-reversed stimuli were represented as different spatial patterns of activation with only some overlap. These activation patterns are presumably of similar spatial extent for natural and backward versions of stimuli, which would explain why activations compared to resting condition were not different.

Thus, differences in results with respect to different reference conditions (stimulus vs. resting condition or stimulus vs. stimulus) in the amygdala and insula as compared to, and in contrast to, AC might reflect differences between sensory cortex maps and brain areas relying on neuronal networks with less mapping of highly processed sensory input (amygdala, insula).

**Lateralizations**

The pattern of lateralized activation of forebrain structures emerging from this study of emotional human specific vocalizations may look peculiar but can be conceptualized in the light of current evidence on lateralized functions of the AC, amygdala, and insular cortex. From an acoustic point of view, laughing and crying as stimuli used here both consist of a number of complex segments separated by pauses, which as a sequence produce the percept of laughter and crying. Single segments do not produce this percept. The left AC has been implicated in the specialized processing of temporal features of sounds and particularly in the high temporal resolution of fast changes (Zatorre, Belin, & Penhune, 2002). We have recently extended this temporal concept to the analysis of sequences of sounds by showing that some regions of the left AC are specifically responsive in a streaming paradigm (Deike, Gaschler-Markefski, Brechmann, & Scheich, 2004). Alternating distinct sounds can be perceptually split into two separate sequences of intrinsically similar sounds, one of which can be followed at a given time. Thus, the task was to sequentially bind similar sounds disregarding the immediate sequence of sounds. In this way, pushing demands on sequential analysis to an extreme, it seems evident that the left AC is highly specialized for sequential analysis. Taken together with a high time resolution for segments, this would explain its usual dominance in speech analysis and also its sensitivity for the present contrast of vocalizations and time-reversed vocalizations. Vocalizations and their reversals are distinct due to both the sequence aspect of segments and the fast dynamics of spectrotemporal changes in the single segments, which are presumably reflected in the left-lateralized direct contrast of activation with forward and backward vocalizations.

Possible anatomical and functional bases of left AC specialization for temporal processing and species-specific vocalizations are provided by monkey and human studies, of which findings can be combined due to their similar anatomical organization of, at least, the primary AC (Wallace, Johnston, & Palmer, 2002). Wang and Kadia (2001) and Wang et al. (1995) detected sub-units of left primary AC neurons in monkeys that responded selectively to natural species-specific vocalizations. Rapid acoustic transients (25 msec) that are typical for complex sounds seem to be represented by discharge rates of primary AC neurons in marmoset monkeys of which the majority exhibited a response asymmetry either in favor of ramped or damped sinuoids (Lu, Liang, & Wang, 2001). In humans, some anatomical data might be interpreted in favor of an improved temporal processing in the left AC: The width of individual cortical columns and the intercolumn distance are larger as are the cell size of magnopyramidal cells and the level of myelination; the number of interconnected columns is smaller (for a review, see Hutslers & Galuske, 2003). Recently, Poremba et al. (2004) demonstrated in rhesus monkeys that intercommissural interactions may be one mechanism by which left lateralization of AC activity in response to conspecific vocalizations is realized. If the two hemispheres were separated, no lateralization of AC activity was observed, indicating that in the intact animal the right hemisphere is suppressed by the left.

Our hypothesis of stronger amygdala activation in response to original rather than time-reversed laughing and crying was corroborated by the present results, however, only in the left amygdala. The involvement of predominantly the left amygdala may be seen in connection with its known engagement in processing emotional auditory stimuli, and involvement of the dominant left AC by direct input from the ipsilateral AC. It should be noted that the exclusively left difference of activation was found for the difference of forward and backward presentations of stimuli. Therefore, this result is not in contrast to our previous study in which only the forward versions of the same stimuli embedded in different tasks led to right-dominant amygdala activation (Sander & Scheich, 2001). Thus, the introduction of and comparison with the backward versions is an independent approach. In addition, the presentation of the emotionally different and unnatural stimuli may have created a context for the forward stimuli that was different from the previous study.

Considering left AC and left amygdala involvement, one would have expected that the left insula would also form part of the hemispheric system processing various aspects of the stimuli. However, the current hypothesis about the role of the right insula in interoception and self-awareness could explain our results on dominant right insula activation by all four stimuli (Craig, 2002, 2004). Craig (2002) argued, on the basis of numerous evidences on the connection between interoception of body states and emotion, that the right insula forms a
metarepresentation of these states and, consequently, a basis for emotional self-awareness. Thus, considering the fact that all our stimuli had emotional quality and that in the context of the whole experiment the backward versions due to their strangeness might have created meta-evaluations, the finding seems compatible with these properties.

In contrast to dichotic listening studies in humans but in line with animal studies and human speech studies, the present results indicate that species-specific human vocalizations like laughing and crying predominantly engage the left AC. This, however, does not imply that all nonspeech vocalizations would in general produce the same results because our results were obtained with segmentally structured vocalizations.

METHODS

Subjects

Fourteen women and six men (age 25.6 ± 5.1 years, mean ± SD), after giving written informed consent, participated in this study, which was approved by the ethical committee of the University of Magdeburg. All subjects had normal hearing and were right-handed (Edinburgh Handedness Inventory) (Oldfield, 1971).

Stimuli

Recordings of continuous laughing and crying produced by professional actors (cf. Sander & Scheich, 2001) were alternately presented in an fMRI block design. Stimulus blocks consisted of either “laughing” or “crying” played in original form (forward) or in time-reversed manner (backward).

Task and Procedure

Using a detection task of occasionally introduced upward shifts of spectral content in each stimulus, we directed the subjects’ attention away from the emotional content and acoustical variation of laughing and crying. Spectral frequencies were increased artificially by 30% for 2 sec (CoolEditPro, Syntrillium Software Corporation, Phoenix, AZ). This was randomly done three to four times per block, resulting in a total of 42 targets. A response had to occur during 3 sec from the beginning of the pitch shifts; otherwise it was not registered as a correct response.

Stimulus presentation was counterbalanced for emotional quality (laughing, crying), gender of voice (female, male), and sequence of presentation mode (forward followed by backward stimulus presentations or vice versa). The experimental session consisted of two baseline blocks at the beginning, 12 stimulus blocks, three for each stimulus, and 12 reference blocks of 45 sec each, resulting in a total duration of 19 min 30 sec. Each stimulus block was followed by a reference block without stimulus presentation, referred to as resting condition. Stimuli were presented binurally with an individually adjusted, comfortable listening level (between 68 and 75 dBA SPL) under computer control.

Immediately before and after scanning, subjects filled out a mood questionnaire to account for possible interactions with the perception of laughing and crying, forward and backward, during the experiment (Mehrdimensionaler Befindlichkeitsfragebogen) (Steyer, Schwenkmezger, Notz, & Eid, 1997). Finally, subjects were debriefed on the pleasantness, emotionality, intensity, and naturalness of each stimulus. Ratings were given on a 7-point scale ranging from 1, the negative pole (unpleasant, etc.), to 7, the positive pole.

In the scanner, subjects’ heads were immobilized with a vacuum cushion with attached ear muffs containing the MRI-compatible electrodynamic headphones (Bamburg et al., 1998).

Data Acquisition

Subjects were scanned in a Bruker (Ettlingen, Germany) Biospec 3T/60 cm system equipped with a quadrupolar head coil and an asymmetrical gradient system (30 mT/m). Volumes of four contiguous axial slices, 8 mm thick and oriented parallel to the Sylvian fissure (Figure 6A), were recorded using a modified noise-reduced FLASH sequence for functional imaging (58 dBA SPL) (Sander, Brechmann, & Scheich, 2003) with interleaved acquisition for each phase-encoding step and with the following parameters: TR = 177.19 msec, TE = 31.65 msec, flip angle = 15°, matrix 64 × 52, FOV 18 × 18 cm². To optimize temporal resolution of the measurement, functional images were recorded by a keyhole procedure. For image formation, highly phase-encoded echoes contribute mostly to object limits and size, whereas low phase-encoding steps are dominantly responsive for object and image contrast and their temporal change. The functional contrast is therefore most dominant in the least phase-encoded echoes. The keyhole scheme reduces the update rate of highly phase-encoded echoes, whereas for image reconstruction, missing phase-encoding steps are taken from a completely phase-encoded reference image that itself is updated at a lower rate. The frequency with which the raw-data matrix is completely actualized defines the keyhole block size. Here, the following specifications were used: keyhole block size = 4, keyhole factor (number of regularly phase-encoding steps divided by the number of completely phase-encoding steps) = 25 / 52 = 0.481, completion position = 3. The choice of the position of the reference image ensures that there is no confusion between image recordings under different neuronal states (Zaitsev, Zilles, & Shah, 2001; Gao et al., 1996; van Vaals et al., 1993).

Eight images were collected during each block, resulting in a total of 208 images for the measured volume.
Directly following fMRI, anatomical images were recorded with the same orientation as the functional images and with high T1 contrast (MDEFT).

Data Analysis

Using the AIR package, functional data were checked for motion (Woods, Grafton, Holmes, Cherry, & Mazziotta, 1998; Woods, Grafton, Watson, Sicotte, & Mazziota, 1998). The criterion for data exclusion was continuous head movements exceeding 1° in one of the three angles or a translation of more than half a voxel in any direction. None of the data sets had to be excluded according to this criterion. Second, functional images were checked for short movement artifacts (due to swallowing, etc.) (see Brechmann, Baumgart, & Scheich, 2002). The mean gray value of all voxels in the region of AC, individually defined across all repetitions, was analyzed because movements can lead to large deviations from mean gray value. Functional images of which gray values deviated more than 2.5% from the mean gray value were excluded from further analysis. A minimum of 20 images has to be left for each condition to perform cross-correlations with an alpha error of 5% and a corresponding beta error of ≤30% for an effect value of 0.8 as recommended for small sample sizes, that is, for small numbers of pairs of comparison (stimulus vs. reference) (Gaschler-Markefski et al., 1997). Therefore, a significance level of 5% error probability was chosen. None of the data sets fell below the limit of 20 images. Thereafter, the matrix size of functional images was increased to 128 × 128 by pixel replication followed by spatial smoothing for reduction of in-plane noise with a Gaussian filter (full width at half maximum = 2.8 mm, kernel width = 7 mm). After mean intensity adjustment of each slice, data sets were corrected for in-plane motion between successive images using the three parameter in-plane rigid body model of BrainVoyager. Functional activation was analyzed by correlation analysis to obtain statistical parametric maps (Gaschler-Markefski et al., 1997; Bandettini, Jesmanowicz, Wong, & Hyde, 1993). A trapezoid function, roughly modeling the expected BOLD response, served as correlation vector with the time series of gray value change of each voxel. Because the development of the BOLD response and its return to baseline takes a few seconds, the first image of each stimulus and reference block was set to half maximum. Images of reference blocks were encoded with minimum value, and images of stimulus blocks with maximum value. Functional data were analyzed with the custom-made software package KHORFu (Gaschler et al., 1996) using cluster pixel analysis of cluster size of eight pixels (in-plane) to exclude false single-point activation. Regions of interest were defined in the left and right hemisphere of each subject for the amygdala complex including some periamygdaloid surrounding the insula and the four different functional-anatomical territories of AC as previously defined (Figure 6B) (Brechmann et al., 2002).

In each subject, significantly activated voxels (p < .05) were attributed to each of the regions of interest (ROIs). Neural activity was determined by contrasting each stimulus versus resting condition as well as by contrasting forward or backward stimulus presentations versus resting condition. Second, neural activity was determined by directly contrasting laughing versus backward laughing, crying versus backward crying, and forward versus backward stimulus presentations. Further analyses were based on the number of activated voxels, IWVs and signal intensity changes in a given ROI as measures of neural activity. IWVs are the product of the number of activated voxels and their mean change in signal intensity (Woldorff et al., 1999). To compute the signal intensity change, the first image of each stimulus and reference block was excluded from data analysis to compensate for the delay of hemodynamic responses. The correlation vector for this computation was therefore similar to a c-box function.

Differences in neural activity due to the factors hemisphere (left, right), stimulus (laughing, crying, backward laughing, and backward crying), and group (D and ND) were analyzed using an analysis of variance with repeated measurements. This was done separately for each of the brain structures and each auditory territory. Post hoc tests were t tests. The level of significance used throughout was 5% error probability, Greenhouse–Geisser epsilon corrected or Bonferroni adjusted for multiple comparisons if necessary. Analyses of (1) target-detection rate, (2) scores of the three scales of the mood questionnaire (see Results), and (3) subjective stimulus ratings were performed using the Wilcoxon matched-pairs test; for the latter, the diagram shows arithmetic means and standard errors of the mean (SEM). The analysis of target detection rate was based on 19 data sets. Due to technical recording failure, one data set was missing.

Acknowledgments

The present study was supported in part by the Deutsche Forschungsgemeinschaft (SFB 426). We thank Dr. T. Kaulisch for methodical support.

The data reported in this experiment have been deposited with the fMRI Data Center archive (www.fmridc.org). The accession number is 2-2005-118PK.

Reprint requests should be sent to Kerstin Sander, Leibniz Institute for Neurobiology, Brennekestrasse 6, 39118 Magdeburg, Germany, or via e-mail: sander@ifn-magdeburg.de.

REFERENCES


