



Published in final edited form as:

Brain Cogn. 2018 July ; 124: 29–36. doi:10.1016/j.bandc.2018.04.001.

Selective Impairments in Components of Affective Prosody In Neurologically Impaired Individuals

Amy Wright¹, Sadhvi Saxena¹, Shannon M. Sheppard¹, and Argye E. Hillis^{1,2,3}

¹Johns Hopkins University School of Medicine, Department of Neurology

²Johns Hopkins University School of Medicine, Department of Physical and Medicine & Rehabilitation

³Johns Hopkins University Department of Cognitive Science

Abstract

The intent and feelings of the speaker are often conveyed less by what they say than by how they say it, in terms of the affective prosody – modulations in pitch, loudness, rate, and rhythm of the speech to convey emotion. Here we propose a cognitive architecture of the perceptual, cognitive, and motor processes underlying recognition and generation of affective prosody. We developed the architecture on the basis of the computational demands of the task, and obtained evidence for various components by identifying neurologically impaired patients with relatively specific deficits in one component. We report analysis of performance across tasks of recognizing and producing affective prosody by four patients (three with right hemisphere stroke and one with frontotemporal dementia). Their distinct patterns of performance across tasks and quality of their abnormal performance provides preliminary evidence that some of the components of the proposed architecture can be selectively impaired by focal brain damage.

Keywords

prosody; emotions; social communication; stroke

Here we propose a preliminary architecture of the perceptual, cognitive, and motor processes underlying recognition and production of affective prosody (variations in pitch, volume, rate, and rhythm of speech to convey emotion in verbal language) on the basis of the computational demands of the task. We obtained support for this architecture through analysis of performance across affective prosody tasks by neurologically impaired participants with distinct impairments in the proposed cognitive architecture. The patterns of performance of four individual cases with relatively selective deficits in affective prosody provide preliminary support for the hypothesis that at least some of the perceptual and cognitive processes underlying affective prosody recognition and production are dissociable.

Corresponding Author: Argye E. Hillis, MD, Department of Neurology, Phipps 446, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA, argye@jhmi.edu, Phone: +1 410-614-2381, Fax: +1 410-614-9807.

Conflicts of Interest

The authors have no conflicts of interest to disclose.

Affective prosody refers to variations in pitch, loudness, rate, and rhythm (pauses, stress, duration of components) of speech to convey emotions. Many studies have shown that recognition and/or production of affective prosody can be impaired by focal (or diffuse) brain damage (Pell, 2002; Pell, 2006), most notably due to right hemisphere stroke (Ross & Mesulam, 1979; Ross & Monnot, 2008; Tippett & Ross, 2015), frontotemporal dementia (Dara et al., 2013; Rankin, Kramer, & Miller, 2005; Rankin et al., 2009), Parkinson's disease (Péron et al., 2015), schizophrenia (Dondaine et al., 2014), and other neurological diseases (Bais, Hoekert, Links, Knegtering, & Aleman, 2010; Kipps, Duggins, McCusker, & Calder, 2007). Several studies have demonstrated dissociations between recognition and production of affective prosody impairment, and some have shown that repetition or mimicking of affective prosody can be disproportionately impaired or spared (Ross & Monnot, 2008). However, few studies have attempted to identify distinct perceptual, cognitive, and motor mechanisms that underlie affective prosody recognition and production [but see (Bowers, Bauer, & Heilman, 1993)].

A number of investigators have proposed neural systems underlying either recognition or production of affective prosody. For example, a neurobiological model of emotional information processing has recently been proposed (Brück, Kreifelts, & Wildgruber, 2011), on the basis of a review of fMRI studies and lesion studies of prosody. In this model, affective prosody is initially processed in the thalamus, which connects to both an explicit pathway and an implicit pathway. The explicit pathway includes extraction of acoustic cues in the middle part of superior temporal gyrus (mSTG), integration of the acoustic cues into a single percept in posterior superior temporal gyrus (pSTG), and evaluation of the percept in dorsolateral prefrontal cortex and orbitofrontal cortex. The implicit pathway projects directly to the amygdala, insula, nucleus accumbens, and medial frontal cortex to induce an emotional response. The authors add that there is evidence, primarily from Parkinson's disease and fMRI that the basal ganglia also have a role in recognition of affective prosody, but the role may be a more general one of processing timing of acoustic information or working memory. A neurobiological model that shares some features is a three stage model proposed by Schirmer and Kotz (Schirmer & Kotz, 2006). This model includes: (1) acoustic analysis in bilateral auditory processing areas; (2) integration of emotionally significant acoustic information into an emotional 'gestalt' along the ventral stream from STG to anterior superior temporal sulcus (STS), which seems to be lateralized to the right hemisphere; and (3) higher-order cognitive processes that yield explicit evaluative judgments of the derived emotional gestalt mediated by the right inferior gyrus and orbitofrontal cortex. Reviews of fMRI studies of affective prosody comprehension have found support for a right-lateralized network including the STG and STS, as well as inferior frontal gyrus (Kotz, Meyer, & Paulmann, 2006; Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006). Although the right hemisphere may have a dominant role in affective prosody, there is clearly a role of the left hemisphere in processing prosody as well, as shown by fMRI studies (Kotz, Meyer, Alter, Besson, von Cramon, & Friederici, 2003; Wildgruber et al., 2006) and lesion studies in which left hemisphere strokes result in impaired production and/or recognition of prosody in speech (Baum & Pell, 1997 & 1999; Cancelliere, & Kertesz, 1990; Schlanger, Schlanger & Gerstman, 1976). Furthermore, integration of information from the two hemispheres, via the corpus callosum, is essential for accurate processing of affective prosody (Paul, Van

Lancker-Sidtis, Schieffer, Dietrich & Brown, 2003). Previous investigations have also provided evidence that elements of prosody perception are differently lateralized, with pitch processed mostly in the right hemisphere, while duration and intensity are processed mainly in the left hemisphere (Van Lancker & Sidtis, 1992; Zatorre, Belin, & Penhune, 2002). In all of the studies that support these neurobiological models, there has been no clear evidence that emotional valence or type of emotion significantly influences the areas that are engaged.

Neurobiological models of *generation* of affective prosody during speech production have implicated a right-lateralized network (Riecker, Wildgruber, Dogil, Grodd, & Ackermann, 2002), bilateral perisylvian network (Aziz-Zadeh, Sheng, & Gheyntchi, 2010), or a critical role of the basal ganglia (Cancelliere and Kertesz, 1990; Van Lancker & Sidtis, 1992). A recent study found support for all three of these networks/regions in different aspects of prosody generation (Pichon & Kell, 2013). In this fMRI study, during the preparatory phases of prosody generation, there was increased ipsilateral connectivity between right ventral and dorsal striatum and between the striatum and the anterior STG, temporal pole, and right anterior insula. Additionally, there was increased connectivity between right dorsal striatum and right inferior frontal gyrus (Brodmann area 44), and left orbitofrontal cortex and inferior temporal gyrus. During the execution phase, connectivity increased between the right ventral striatum and dorsal striatum, and between dorsal striatum and bilateral inferior frontal gyrus and right STG, but also between right dorsal striatum and right anterior hippocampus and bilateral amygdala.

While these neurobiological models are useful for understanding areas of the brain engaged in affective prosody recognition and/or production, they are underspecified in terms of the cognitive processes that underlie the ubiquitous and socially imperative tasks of recognizing and conveying emotion through prosody. How is it that everyone in the room is able to recognize when a speaker is angry, sorrowful, afraid, or delighted, not by what they say, but how they say it? Likewise, how is it that we all convey our own emotions not so much by what we say, but how we say it, in such a way that the emotion is universally understood (even if the listener does not share the language)?

To develop a cognitive architecture of the representations and processes underlying recognition and expression of affective prosody, we can start by considering the computational demands of recognizing and expressing emotions through changes in the acoustic features of speech. To recognize emotion in another's speech, it is first essential to parse and analyze the paralinguistic features of the utterance. This acoustic analysis requires recognition of differences in pitch (frequency), loudness (intensity), rate, and rhythm (stress, pauses, and duration of various segments). Then, it is necessary to match a set of acoustic features with an emotion. To do this, we need access to an abstract representation of what "angry", "sad", "happy" and so on sound like. These abstract representations of acoustic characteristics that convey emotion (ARACCE) are shared by speakers of a language or culture. The ARACCE (whether stored or computed on-line) would specify the acoustic features (e.g., low frequency, high intensity, rapid rate and their interactions) of each emotion (e.g., anger), and allow access to the semantic representation, or meaning, of that emotion. Thus, the ARACCE is comparable to the lexical orthographic representation for reading and spelling, in that it mediates between the semantic representation and input or

output processes. The semantic representation would specify an autonomic response (i.e., the so-called “fight or flight” or sympathetic response or parasympathetic response) as well as the meaning and valence, and would overlap with semantic representations of other emotions. For example, disappointed and heartbroken share aspects of meaning, but vary in valence. To express an emotion (e.g. anger) through prosody, one would need to access the abstract representation of the acoustic features of anger (the ARACCE) from the semantic representation of anger. Then, one would need to convert the ARACCE to motor programs for producing these acoustic features (e.g. changes in length and tension of the vocal folds by constriction or relaxation of cricothyroid and other laryngeal and respiratory muscles, changes in rate of movement of the lips, tongue, jaw, palate, and so on). Finally, one would need to *implement* these motor programs during the complex act of speaking. These proposed mechanisms are schematically represented in Figure 1. Although we have discussed these mechanisms as though they are activated serially, it is likely that some operate in parallel during listening or production of speech. Also, there may be feedforward or feedback interactions between levels that are not depicted here.

In this study, we provide empirical evidence for some of the proposed levels of processing, by showing that they can be relatively selectively disrupted by brain damage. Participants with acute right hemisphere stroke or frontotemporal dementia (FTD) were tested on a battery of tests designed to assess processing at each proposed level. We selected these two populations because impaired prosody is common after both right hemisphere stroke (Dara, Bang, Gottesman, & Hillis, 2014; Ross & Mesulam, 1979) and FTD (Dara et al., 2013; Rankin et al., 2009); Phillips, Sunderland-Foggio, Wright, & Hillis, 2017). Unlike individuals with autism or schizophrenia, it is possible to identify the region of infarct or focal atrophy associated with the impaired level of processing.

In this ongoing study, many participants showed no impairment at any level and others had deficits at several levels of processing or insufficient testing to determine the level(s) of processing that were impaired. Nevertheless, we identified four participants whose performance across tasks and acoustic characteristics of speech can be accounted for by assuming selective disruption of a single level.

Methods

Participants

A total of four neurologically impaired individuals were included in this case series, including one with FTD and three participants with acute, right hemisphere stroke. All participants were recruited from the Johns Hopkins Hospital stroke service or the cognitive disorders clinic; all were right-handed. They were selected from a total of nine patients with FTD and 22 patients with acute ischemic right hemisphere stroke who completed the Affective Prosody Battery. Additionally, 60 healthy, age- and education-matched neurotypical controls were tested on one or more of the tests (see numbers for each test in Table 1). Exclusion criteria for both neurologically-impaired participants and controls included: diminished level of consciousness or sedation, lack of premorbid proficiency in English, and uncorrected hearing or visual loss. For neurologically impaired participants, an

inclusion criterion was relatively selective impairment in one level of processing in affective prosody. All participants provided informed consent for the study.

Case 1 is a 61 year old right handed man who was initially tested within 48 hours of an acute ischemic stroke. At the time of his stroke he held a high position in business, and he had a graduate education. MRI with diffusion weight imaging and arterial spin labeling perfusion imaging (completed the same day of testing) showed acute infarct in the right anterior temporal cortex, amygdala, and anterior insula (Figure 2). The stroke was thought to be cardioembolic, and he showed spontaneous reperfusion of the affected area.

On neurological examination, he had fluent, grammatical speech with normal prosody. He scored in a non-depressed range on the Brief Patient Health Questionnaire-9 (PHQ-9). On the Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) he scored 25/30. On the MMSE he lost points on calculation; he made four errors on subtracting serial sevens. He also lost 1 point on the recall task where he could only recall 2 out of 3 words after intervening tasks. Additionally, he had very mild stimulus-centered neglect.

Case 2 is a 60 year old woman with a past medical history significant for hypertension, hyperlipidemia, and mitral valve prolapse who developed acute onset left-sided hemiplegia while ambulating in the bathroom. She fell and was not able to get up until she was found hours later to have left sided weakness and a left sided facial droop. In the Emergency Department, head CT showed hypodensity in the territory of the right middle cerebral artery. MRI performed the next day, the same day as behavioral testing, confirmed acute infarct in right insula and superior temporal cortex (Figure 3).

On examination, she was alert and oriented to person, place, time, and situation, and she had normal recent and remote memory. Her MMSE score of 29/30 was within normal limits. She did not have aphasia, and demonstrated normal fluency and comprehension. She was able to follow one and two-step commands, and repeat simple and complex sentences. She did not have right/left confusion. Her calculation abilities were intact (e.g. able to state number of quarters in \$1.75). She did not have apraxia. She did have mild hemispatial neglect.

Case 3 is a 56 year old male manager who was diagnosed with frontotemporal dementia (FTD) about one year prior to testing. His MMSE score was normal at 29/30. His scores were normal on tests of executive functioning. However, his wife reported obsessive compulsive behaviors, apathy, and disinhibition. She also reported that he had marked decline in empathy, initiative, emotional responsiveness and energy. His speech was fluent and grammatical, and his comprehension was normal when following complex commands. He had normal forward and backward digit spans. He did not have paranoia or hallucinations. He also scored in the non-depressed range on the PHQ-9, and had a normal neurological exam.

Case 4 is a 56 year old woman who presented to the Emergency Department with left sensory loss, weakness on the left arm and leg, left facial droop and hemispatial neglect. However, her symptoms improved, and she had only residual left sided numbness and heaviness and flat tone of voice. MRI showed a small area of acute infarct in the right insula

and the posterior frontal periventricular white matter and a larger area of hypoperfusion including right STG (Figure 4).

On neurological examination following mechanical thrombectomy, she was awake, alert, oriented, and appropriately interactive. Her naming, comprehension, and repetition were intact for content. She did not extinguish left sided visual or tactile stimuli to double simultaneous stimulation, but she did have motor extinction: on motor tasks like finger tapping or arm abduction she stopped using her left limb when also using right limb (Hillis et al., 2006). She had no cranial nerve deficits except a mild left lower facial droop. Her motor exam was notable for mild weakness in the left leg.

Affective Prosody Battery

Tests were developed to assess each of the proposed levels of processing. Several subtests are modeled after the Aprosodia Battery (Ross & Monnot, 2008), including sentence repetition (and evaluation of the coefficient of variation in fundamental frequency to assess prosodic production) and recognition of emotions in sentences. However, we used novel sentences and additional subtests to disambiguate deficits at various levels. The participants were trained in each task with examples, and demonstrated comprehension of acoustic features (e.g., slow rate, monotone, rising pitch) through correct identification of these features (from eight choices) to match a simple auditory stimulus (e.g., pure tone or series of tones) in comparison with a model stimulus (a contrasting pure tone or series of tones). They completed the following sentence by pointing to a label of the feature: “Compared to the first stimulus/set, the second one was ____ ” To ensure that they understood the concept of “rate”, they listened to a series of tones for each stimulus (one rapid, one slow), and indicated the rate of the second, relative to the first. They could repeat this “screening” test until they achieved 90% correct. None of the participants required more than 2 repetitions.

The acoustic features corresponding to each emotion were identified in studies of healthy participants (Banse & Scherer, 1996; Johnstone & Scherer, 1999; Scherer, Johnstone, & Klasmeyer, 2003). Impairment on matching tasks (described below) were identified if performance was 2 standard deviations from the mean for the neurotypical controls. Deficits in production tasks were measured by mean Coefficient of Variation in Fundamental Frequency (F_0 -CV%) across items in the tasks. All responses were recorded, and each individual sentence is spliced. Their utterances then undergo acoustical analysis using Praat software (Boersma et al., 2018) to measure variation in fundamental frequency (F_0) over time across each response. F_0 -CV% was selected because this measure correlated with listener judgments of intactness of affective prosody in sentence repetition (Basilakos et al., 2017). Furthermore, the ability to manipulate F_0 to produce intonation is the most “salient acoustical feature underlying affective prosody in speakers of non-tone languages, such as English” (Ross & Monnot, 1980; p. 54). The intraclass correlation coefficient (ICC) for the calculation of F_0 -CV% was 0.982 (95% CI = 0.945 – 0.994), which demonstrated high inter-rater reliability. Impairment was identified by 2 standard deviations below the mean for the controls.

Recognition of prosodic features was evaluated by administering the following task:

1. Identification of acoustic features in sentences. Participants listened to audio files of 24 neutral content sentences (e.g. “She will be going to the new school”) spoken by a female speaker with one of six emotions (sad, angry, disgusted, fearful, surprised, or happy), and pointed to two acoustic features that corresponded to each sentence production (from nine possible: high pitch; low pitch; monotone; rising pitch; declining pitch; high volume; low volume, fast rate; slow rate). Many of the sentences had at least three correct features; but the participant could get credit for any two correct features (48 possible points)

Access to the ARACCE (from auditory input) was evaluated with the following tasks:

1. Identification of features to match emotion words. Participants were asked to select at least two features (from nine: high pitch; low pitch; monotone; rising pitch; declining pitch; high volume; low volume, fast rate; slow rate) associated with each of five emotion words (spoken and printed on a card). Features were identified from studies of healthy controls (Banse & Scherer, 1996; Scherer et al., 2003).
2. Identification of emotions from sentences. Participants listened to 24 semantically neutral sentences produced with emotional intonation and selected the corresponding emotion (sad, angry, disgusted, fearful, surprised, or happy) from five choices.

Access to Semantics of Emotions was evaluated with the following tasks:

1. Matching synonyms: Participants were asked to select one word from two choices, that was closest in meaning to a given emotion (e.g., joy: pleasure or affection) (n=24). To determine if impairment was specific to emotions, they also were asked to select one non-emotion abstract word from two choices, closest in meaning to a non-emotion abstract word (e.g., reality: fact or independence) (n=24). Emotions and other abstract words were matched in word frequency.
2. Matching emotions to situations: Participants were asked to select one emotion from six choices (sad, angry, disgusted, fearful, surprised, or happy) to match a given situation (e.g., You are cheated out of money that you did not really need).
3. Matching emotion labels to pictures of faces depicting each emotion: Participants were asked to select one emotion from six choices (angry/disgusted, sad, fearful, happy, surprised) to match a face showing that emotion. The stimuli for this task were selected from a set of perceptually validated pictures (Pell & Leonard, 2005). Note that poor performance on this task could be due to impaired understanding of the meaning of the emotion or impaired recognition of facial expressions of emotions. However, normal performance on the task indicates spared understanding of the meaning of the emotion and spared recognition of facial expressions of emotions. We combined angry/disgusted (scored as correct if they chose either emotion for the angry or disgusted faces, because our healthy controls were poor at distinguishing angry and disgusted faces.

Access to the ARACCE for Production was evaluated by administering the following tasks:

1. Sentence Reading with a Given Emotion. Participants were asked to read aloud 24 semantically neutral sentences with a specified emotion (four trials for each emotion: sad, angry, disgusted, fearful, surprised, or happy). Sentences were typed in 14 point font on white paper.
2. Sentence Reading with Cues: Participants were asked to again read aloud the 24 sentences with a given emotion (from above task), this time given the correct prosodic features corresponding to the emotion (determined by acoustic analysis of neurotypical speakers conveying each emotion (Banse & Scherer, 1996; Scherer et al., 2003). Sentences and the cues were typed in 14 point font on white paper. The cues were typed to the right of the sentence. The prosodic cues were the same acoustic features used in the recognition of prosodic features tasks (e.g. high pitched, fast) described above. Significantly impaired (low) F_0 -CV% in reading sentences without cues, but significantly higher with cues (the prosodic features) was considered evidence of impaired access to the ARACCE for production.
3. Sentence Repetition: Participants were asked to listen to 24 sentences, and repeat each one with the same emotion as the speaker, but were not told the emotion. Significantly higher F_0 -CV% in repetition compared to reading was considered evidence for impaired access to ARACCE for production.

Motor planning and implementation were evaluated with the above two tasks. If the participant showed low F_0 -CV% in production with cues and in repetition, which was not significantly higher than F_0 -CV% in reading the sentences without cues, they were considered to have impaired motor planning or implementation. Note that these tasks did not differentiate motor planning from implementation.

Results

For this study, we selected four neurologically impaired participants who had a pattern of performance that could be accounted for by assuming relatively selective damage to one or two components of the proposed perceptual, cognitive, and motor mechanism underlying affective prosody. The details of each case are described below.

Impaired Extraction of Prosodic Features in Speech

Performance on assessments of prosody by Case 1, who had an acute stroke involving right anterior temporal and insular cortex, and amygdala, is reported in Table 1. Relative to controls in his age range, he was impaired in identifying acoustic features (e.g. low pitch, low volume) from sentences, although he was able to identify the same features when comparing pairs or series of tones. Thus, he demonstrated intact comprehension of the feature labels in simple tones, but was impaired in identifying the features from speech.

Case 1 was also impaired in identifying the emotion conveyed in sentences (which requires both extraction of the acoustic features and access to the ARACCE). Yet he showed normal performance in matching prosodic features to emotions in speech, which also requires access to the ARACCE. Thus, he showed intact comprehension of the labels of emotions and intact

access to the ARACCE. His performance was also normal on tasks of prosody semantics and expressive prosody. Therefore, he seemed to have a selective deficit in extracting acoustic features from speech.

Scores on assessments of prosody by Case 2, who had an acute stroke involving right superior temporal cortex and insula, are reported in Table 1. Relative to controls in her age range, she was impaired in identifying prosodic features (e.g. low pitch, low volume) from sentences, although she was able to identify the same features when asked to describe the difference between pairs of simple or series of tones. In this way, she demonstrated intact comprehension of the feature labels, but was impaired in identifying the prosodic features from spoken sentences.

Like Case 1, Case 2 was also impaired in identifying the emotion conveyed in sentences (which requires both extraction of the acoustic features and access to the ARACCE). She showed normal performance in matching prosodic features to emotions in speech, which also requires access to the ARACCE. Therefore, she demonstrated intact comprehension of the labels of emotions and intact access to the ARACCE. Her performance was intact on tasks of prosody semantics and expressive prosody. Thus, her impairment in recognizing emotions seemed to be due to a selective impairment in extracting acoustic features from speech.

Cases 1 and 2 had a F_0 -CV% that was more than 1 SD *above* the mean for controls. We believe they were just trying to be dramatic.

Impaired Access to the ARACCE for Recognition of Affective Prosody

Case 3, the individual with early FTD, showed normal performance in recognition of prosodic features on our affective prosody battery, indicating that he understood the names of the features and could detect them in speech. However, when given a particular emotion word (e.g., anger) he could not correctly match acoustic features that are associated with that emotion (e.g., low pitch) from a set of written choices (Table 1). He was also impaired in identifying an emotion to match a spoken sentence. Both of these tasks require access to the ARACCE and the meaning of emotions. He showed spared comprehension of emotions on the semantic tasks. Thus, his impairments in matching acoustic features to emotion and in matching spoken sentences to emotions could be attributed to selective impairment in accessing information about what emotions sound like – the ARACCE – from spoken input.

Case 3 was not impaired in accessing the ARACCE for spoken output. He had intact production of emotions in sentences (normal F_0 -CV% relative to controls) in repetition and reading of sentences, ruling out a problem in accessing the ARACCE for output, as well as a problem in motor planning or implementation. Therefore, his impaired access to the ARACCE for recognition of prosody was a relatively selective deficit.

Impaired access to the ARACCE for generation of affective prosody

Performance of Case 4, who had an acute right fronto-insular stroke, with hypoperfusion in left superior temporal gyrus, is also reported in Table 1. On the affective prosody battery, she was 100% accurate on her first attempt at the screening test, and showed normal

performance on all subtests of prosody recognition, identifying acoustic features associated with each emotion, and semantics of emotion. However, she had significantly reduced F_0 -CV% in reading sentences with specified emotions (Table 1). Importantly, her F_0 -CV% improved to normal when she was given cues (the appropriate acoustic features, such as slow, low pitched, quiet for sad). She also showed normal F_0 -CV% in sentence repetition with instructions to say the sentence with the same emotion as the speaker, indicating that she did not have a substantial motor speech impairment that could account for the reduced F_0 -CV% in sentence reading without cues. Her response to the cues indicates that she had impaired access to ARACCE (the abstract representation of acoustic characteristics that convey emotion) for output.

Discussion

Here we have proposed an architecture of the perceptual, semantic, and motor processes required for recognition and production of affective prosody. We have provided evidence for some of these component mechanisms. This architecture (of largely right hemisphere processes and representations) is very similar to proposed architectures of lexical processing for propositional speech comprehension and production (largely in the left hemisphere; e.g. Hillis, Rapp, Romani & Caramazza, 1990).

Three cases demonstrate that it is possible to identify dissociable impairments in recognition of affective prosody. That is, two had perceptual impairments (Cases 1 & 2), with difficulty identifying acoustic features from speech, which also led to impairment in identifying emotions from speech. The third (Case 3) showed intact perception (identification of acoustic features from speech) but impaired access to the ARACCE, which manifested as difficulty identifying emotions but not acoustic features from speech and difficulty correctly matching an emotion word (e.g., anger) to the correct acoustic features from a set of written choices (e.g., low pitch/high pitch, slow rate/fast rate). These cases indicate that “receptive aprosodia” (Ross & Monnot, 2008) can result from two distinct levels of processing – one being perceptual and the other in accessing or computing and abstract information about what an emotion sounds like.

This study provides a conceptual framework for identifying dissociable impairments in recognition and expression of affective prosody. It may well be modified by subsequent evidence. Furthermore, we did not provide evidence for selective deficits in all of the components. We have identified individuals with impairment in other components, like semantic representations of emotions, but have not been able to demonstrate that the deficit is selective. The impairment is reflected in failing to match names of emotions to faces or phrases describing an emotional event, impaired matching of emotions to synonyms (with spared matching frequency-matched emotions to a synonyms). However, in these cases it is impossible to determine the status of access to the ARACCE, because impaired understanding of the meaning of emotion words also disrupts performance on matching acoustic features to emotions and matching spoken sentences to emotions. In contrast to evaluating lexical-semantics after left hemisphere stroke, which results in similar semantic errors across tasks (Hillis, Rapp, Romani & Caramazza, 1990), all errors on our recognition tasks could be considered “semantic errors” because forced choice responses are limited to

semantically related words (happy, sad, angry, etc.). It is also difficult to evaluate whether production deficits are semantic errors (e.g. sounding sad rather than angry) or other type of error.

Participants with deficits in different components of this architecture had different areas of brain damage (stroke or focal atrophy). However, it is not possible to make claims about the areas or networks essential for any component until a sufficient number of participants are identified with and without disruption of that level of processing to show a statistical association between the prosody deficit and the region of damage.

Nevertheless, the few cases reported here are consistent with the proposal that processes involved in recognition of affective prosody depend on a ventral stream of processing – a sound to meaning network from superior temporal cortex to inferior and anterior temporal cortex – comparable to the ventral stream of processing of language in the left hemisphere (see also Wright et al., 2016 for additional preliminary evidence from stroke and Schirmer & Kotz, 2006 for a similar view based on functional imaging studies). That is, the two individuals with impaired access to acoustic features from speech (Cases 1 and 2) had right superior or anterior temporal lesions (also right anterior insula and anterior basal ganglia in Case 1). This deficit is reminiscent of “pure word deafness” except that it involved the processing of affective prosody. Pure word deafness is most often associated with focal lesions involving the left auditory cortex (mid STG area) that also disrupts callosal fibers from the right auditory cortex leaving most of posterior STG intact. Although Case 1 did not appear to have acute ischemia involving the right auditory cortex, he was examined during the acute phase of his stroke. There was no evidence on perfusion imaging (Fig 1) of an ischemic penumbra that affected the right auditory cortex, but there is a possibility of diaschisis acutely disrupting the right auditory cortex. Case 2, who also had impaired access to acoustic features of emotions from speech, clearly had acute infarct involving right middle STG. The potential “affective” acoustic analyzer in right middle STG may be analogous to the proposed role of left middle STG in propositional language.

We propose that Case 3 had impaired access to an “abstract representations of acoustic characteristics that convey emotion (ARACCE)”. This proposed right hemisphere mechanism is reminiscent of Wernicke’s proposed left hemisphere mechanism in the “ideational area” that is required for word comprehension. It is also analogous to access to a spoken word form (or phonological input lexicon). We are not able to draw conclusions about the localization of this function, as Case 3 had fairly diffuse right greater than left temporal atrophy.

Also of note, the few individuals we have identified with semantic deficits (with or without other deficits) have had bilateral (right greater than left) anterior temporal atrophy. We have not observed this impairment after unilateral stroke. The deficit may be analogous to the more general amodal semantic deficit observed in people with semantic variant primary progressive aphasia (with left greater than right anterior temporal atrophy) (see also Zahn, et al., 2009).

Finally, the individual with a subcortical parietal infarct and hypoperfusion of inferior parietal cortex had selective impairment of accessing the ARACCE for output. This is consistent with the proposal that prosody expression depends on a dorsal stream of processing, from temporal cortex, to inferior parietal cortex, to frontal cortex – a “meaning to motor production” network comparable to a dorsal stream of processing of language in the left hemisphere (Kummerer et al., 2013; Poeppel & Hickok, 2004; Saur & Hartwigsen, 2012). Previous studies have indicated that the area most likely associated with the direct motor programming of affective prosody is right inferior-posterior frontal region, homologue of Broca’s area) (Ross & Monnot, 1980, Wright et al., 2016), also in the dorsal stream.

However, there is evidence from both functional imaging and lesion data that limbic system (including anterior cingulate, anterior insula, amygdala) are also critical for processing emotions (Tippett & Ross, 2015). All of our patients also had damage (acute infarct or atrophy) to the right insula, so it is possible that insular damage contributed to one or more impairments in affective prosody. Further study of patients with and without insular damage, and with and without deficits in prosody processing will yield important information toward clarifying the role of the insula in affective prosody.

We did not include participants with left hemisphere lesions or predominant left hemisphere focal atrophy in this study, because participants with deficits due to left hemisphere damage typically have aphasia and have trouble understanding the tasks. Many previous lesion studies and functional neuroimaging studies of healthy controls have provided evidence that affective prosody depends more on the right hemisphere than left hemisphere.

Limitations of the architecture we have proposed include the fact that each of the proposed mechanisms may depend on a number of distinct sub-processes that might be dissociable. For example, acoustic analysis of prosodic features includes identification of pitch, loudness, rate, and rhythm (duration of various segments, pauses, and stress). It is very possible that brain damage (or more peripheral hearing impairment) might selectively impair processing of a single acoustic feature. Analysis of errors to date has not yet identified such a specific deficit. Similarly, functional imaging studies of affective prosody have not identified differential sites of activation associated with distinct acoustic features. Likewise, performance on the impaired motor planning and implementation tasks could be due to impairment in motor planning/programming or due to dysarthria (a deficit in rate, range of motion, strength, or coordination of the muscles of speech articulation). Analysis of errors may be able to distinguish between these deficits; e.g., consistent errors on the same phonemes are more consistent with dysarthria. Moreover, a careful neurological examination of the strength, range of movement, and rate of movement of the muscles of the jaw, face, lips, tongue, pharynx, larynx, and respiration, as well as apraxia of speech is essential. There also might be cases of “category specific” errors in semantics of emotions (e.g. difficulty understanding positive or negative emotions, or emotions related to fear), although our one patient with a semantic deficit made errors on all emotions.

The preliminary battery we have developed has some weaknesses as well. For example, additional acoustic features, beyond $F_0CV\%$, that correlate with listener judgments of affective prosody, will likely improve sensitivity to impairments in components of prosody

expression. $F_0CV\%$ measures variation in pitch, but as indicated by perceptual studies of neurotypical speech, emotion is also conveyed by variations in loudness, rate, and stress patterns (see Basilakos et al., 2017). Also, the multiple choice format of many of the subtests may reduce the sensitivity (but increases the reliability in scoring). All of the tasks are metalinguistic; it is necessary to determine how performance on these metalinguistic tasks corresponds to affective prosody deficits in more natural environments. Previous studies have shown that reading and repetition differ significantly from spontaneous (or elicited) speech in important motor speech measures (Johns & Darley, 1970; Kempler & Van Lancker, 2002; Van Lancker Sidtis, Cameron, & Sidtis, 2012). Evaluation of prosody in elicited connected speech (e.g. describing emotional events) will be important in future studies.

Despite its limitations, this study has provided preliminary support for distinct perceptual and cognitive processes underlying recognition and production of affective prosody. Future studies are planned to (1) identify more individuals who can be shown to have disrupted or intact mechanisms underlying affective prosody, to extend and confirm these findings and to allow statistical analysis of associations between deficits and lesion sites; (2) refine the assessment of prosody to allow identification of more specific impairments within each processing component, and (3) explore the relationship between impaired performance on these metalinguistic tasks and problems with social interactions and functional communication.

Acknowledgements

This work was supported by the National Institutes of Health (National Institute of Neurological Disorders and Stroke and National Institute on Deafness and Communication Disorders) through award R01 NS4769 and R01 DC05375. The content is solely the responsibility of the authors and does not necessarily represent the views the National Institutes of Health. We are grateful for this support and to the participants in the study, and to the two anonymous reviewers who made helpful and insightful suggestions to improve the paper.

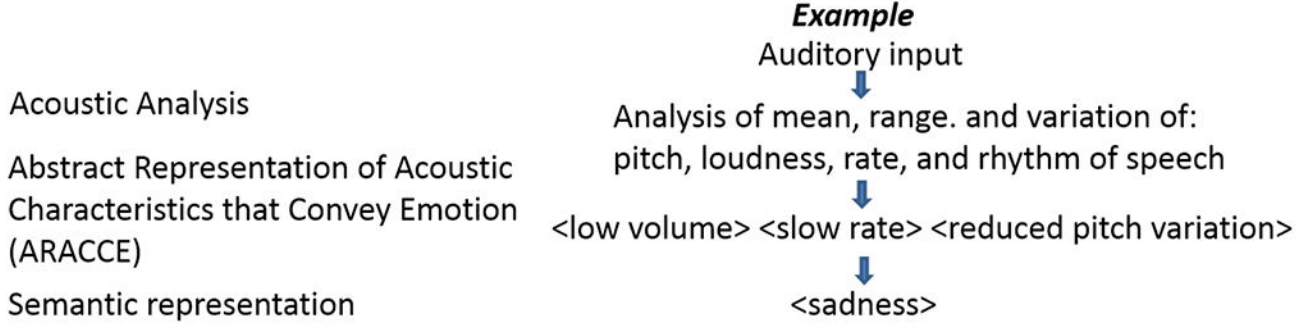
References

- Aziz-Zadeh L, Sheng T, & Gheytauchi A (2010). Common premotor regions for the perception and production of prosody and correlations with empathy and prosodic ability. *PloS One*, 5(1), e8759. [PubMed: 20098696]
- Bais L, Hoekert M, Links TP, Knegtering R, & Aleman A (2010). Brain asymmetry for emotional prosody in klinefelter's syndrome: Causal relations investigated with tms. *Schizophrenia Research*, 117(2), 233–234
- Banse R, & Scherer KR (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70(3), 614. [PubMed: 8851745]
- Basilakos A, Demsky C, Sutherland-Foggio H, Saxena S, Wright A, Tippett DC, . . . Hillis AE (2017). What makes right hemisphere stroke patients sound emotionless? [abstract] *Neurology*, S6, 327.
- Baum SR, & Pell MD (1997). Production of affective and linguistic prosody by brain-damaged patients. *Aphasiology*, 11(2), 177–198.
- Baum SR, & Pell MD (1999). The neural bases of prosody: Insights from lesion studies and neuroimaging. *Aphasiology*, 13(8), 581–608.
- Boersma Paul & Weenink David (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.39, retrieved 3 April 2018 from <http://www.praat.org/>
- Bowers D, Bauer RM, & Heilman KM (1993). The nonverbal affect lexicon: Theoretical perspectives from neuropsychological studies of affect perception. *Neuropsychology*, 7(4), 433.

- Brück C, Kreifelts B, & Wildgruber D (2011). Emotional voices in context: A neurobiological model of multimodal affective information processing. *Physics of Life Reviews*, 8(4), 383–403. [PubMed: 22035772]
- Cancelliere AE, & Kertesz A (1990). Lesion localization in acquired deficits of emotional expression and comprehension. *Brain and cognition*, 13(2), 133–147. [PubMed: 1697174]
- Dara C, Kirsch-Darrow L, Ochfeld E, Slenz J, Agranovich A, Vasconcellos-Faria A, . . . Kortte KB (2013). Impaired emotion processing from vocal and facial cues in frontotemporal dementia compared to right hemisphere stroke. *Neurocase*, 19(6), 521–529. [PubMed: 22827701]
- Dara C, Bang J, Gottesman RF, & Hillis AE (2014). Right hemisphere dysfunction is better predicted by emotional prosody impairments as compared to neglect. *Journal of Neurology & Translational Neuroscience*, 2(1), 1037. [PubMed: 24945016]
- Dondaine T, Robert G, Péron J, Grandjean D, Vérin M, Drapier D, & Millet B (2014). Biases in facial and vocal emotion recognition in chronic schizophrenia. *Frontiers in Psychology*, 5 [PubMed: 24478738]
- Folstein MF, Folstein SE, & McHugh PR (1975). “Mini-mental state”: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198. [PubMed: 1202204]
- Hillis AE, Rapp BC, Romani C & Caramazza A (1990). Selective impairment of semantics in lexical processing. *Cognitive Neuropsychology*, 7, 191–244
- Hillis AE, Chang S, Heidler-Gary J, Newhart M, Kleinman JT, Davis C, ... & Ken L (2006). Neural correlates of modality-specific spatial extinction. *Journal of Cognitive Neuroscience*, 18(11), 1889–1898. [PubMed: 17069479]
- Johns DF, & Darley FL (1970). Phonemic variability in apraxia of speech. *Journal of Speech, Language, and Hearing Research*, 13(3), 556–583.
- Johnstone T, & Scherer KR (1999). The effects of emotions on voice quality. *Proceedings of the XIVth International Congress of Phonetic Sciences*, 2029–2032.
- Kaplan E, Goodglass H, Weintraub S, Segal O, & van Loon-Vervoorn A (2001). *Boston naming test*. New York: Pro-Ed.
- Kertesz A (2007). *The western aphasia battery - revised*. New York: Grune & Stratton.
- Kempler D, & Van Lancker D (2002). Effect of speech task on intelligibility in dysarthria: A case study of Parkinson’s disease. *Brain and Language*, 80(3), 449–464. [PubMed: 11896652]
- Kipps C, Duggins A, McCusker E, & Calder A (2007). Disgust and happiness recognition correlate with anteroventral insula and amygdala volume respectively in preclinical huntington’s disease. *Journal of Cognitive Neuroscience*, 19(7), 1206–1217. [PubMed: 17583995]
- Kotz SA, Meyer M, & Paulmann S (2006). Lateralization of emotional prosody in the brain: An overview and synopsis on the impact of study design. *Progress in Brain Research*, 156, 285–294. [PubMed: 17015086]
- Kotz SA, Meyer M, Alter K, Besson M, von Cramon DY, & Friederici AD (2003). On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain and language*, 86(3), 366–376. [PubMed: 12972367]
- Kummerer D, Hartwigsen G, Kellmeyer P, Glauche V, Mader I, Kloppel S, . . . Saur D (2013). Damage to ventral and dorsal language pathways in acute aphasia. *Brain : A Journal of Neurology*, 136(Pt 2), 619–629. doi:10.1093/brain/aws354 [doi] [PubMed: 23378217]
- Paul LK, Van Lancker-Sidtis D, Schieffer B, Dietrich R, & Brown WS (2003). Communicative deficits in agenesis of the corpus callosum: nonliteral language and affective prosody. *Brain and Language*, 85(2), 313–324. [PubMed: 12735947]
- Pell MD (2002). Evaluation of nonverbal emotion in face and voice: Some preliminary findings on a new battery of tests. *Brain and Cognition*, 48(2–3), 499–514. [PubMed: 12030496]
- Pell MD (2006). Cerebral mechanisms for understanding emotional prosody in speech. *Brain and Language*, 96(2), 221–234. [PubMed: 15913754]
- Pell MD, & Leonard CL (2005). Facial expression decoding in early Parkinson’s disease. *Cognitive Brain Research*, 23(2), 327–340. [PubMed: 15820640]

- Péron J, Cekic S, Haegelen C, Sauleau P, Patel S, Drapier D, . . . Grandjean D (2015). Sensory contribution to vocal emotion deficit in parkinson's disease after subthalamic stimulation. *Cortex*, 63, 172–183. [PubMed: 25282055]
- Pichon S, & Kell CA (2013). Affective and sensorimotor components of emotional prosody generation. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 33(4), 1640–1650. doi:10.1523/JNEUROSCI.3530-12.2013 [doi]
- Phillips L, Sunderland-Foggio H, Wright A, & Hillis AE (2017). Impaired recognition of sadness in face or voice correlates with spouses' perception of empathy in frontotemporal dementia. [abstract]. *Neurology*, S4, 184.
- Poeppel D, & Hickok G (2004). Towards a new functional anatomy of language. *Cognition*, 92(1), 1–12. [PubMed: 15037124]
- Rankin KP, Kramer JH, & Miller BL (2005). Patterns of cognitive and emotional empathy in frontotemporal lobar degeneration. *Cognitive and Behavioral Neurology : Official Journal of the Society for Behavioral and Cognitive Neurology*, 18(1), 28–36. doi:00146965-200503000-00004 [pii] [PubMed: 15761274]
- Rankin KP, Salazar A, Gorno-Tempini ML, Sollberger M, Wilson SM, Pavlic D, . . . Miller BL (2009). Detecting sarcasm from paralinguistic cues: Anatomic and cognitive correlates in neurodegenerative disease. *NeuroImage*, 47(4), 2005–2015. [PubMed: 19501175]
- Riecker A, Wildgruber D, Dogil G, Grodd W, & Ackermann H (2002). Hemispheric lateralization effects of rhythm implementation during syllable repetitions: An fMRI study. *NeuroImage*, 16(1), 169–176. [PubMed: 11969327]
- Ross ED, & Mesulam M (1979). Dominant language functions of the right hemisphere?: Prosody and emotional gesturing. *Archives of Neurology*, 36(3), 144–148. [PubMed: 435134]
- Ross ED, & Monnot M (2008). Neurology of affective prosody and its functional–anatomic organization in right hemisphere. *Brain and Language*, 104(1), 51–74. [PubMed: 17537499]
- Saur D, & Hartwigsen G (2012). Neurobiology of language recovery after stroke: Lessons from neuroimaging studies. *Archives of Physical Medicine and Rehabilitation*, 93(1), S15–S25. [PubMed: 22202187]
- Scherer KR, Johnstone T, & Klasmeyer G (2003). Vocal expression of emotion. *Handbook of Affective Sciences*, , 433–456.
- Schlanger BB, Schlanger P, & Gerstman LJ (1976). The perception of emotionally toned sentences by right hemisphere-damaged and aphasic subjects. *Brain and Language*, 3(3), 396–403. [PubMed: 949593]
- Schirmer A, & Kotz SA (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, 10(1), 24–30. [PubMed: 16321562]
- Tippett DC, & Ross E (2015). Prosody and the aprosodias In Hillis AE (Ed.), *The handbook of adult language disorders* (2nd ed, pp. 518–529). New York: Psychology
- Van Lancker Sidtis D, Cameron K, & Sidtis JJ (2012). Dramatic effects of speech task on motor and linguistic planning in severely dysfluent parkinsonian speech. *Clinical Linguistics & Phonetics*, 26(8), 695–711. Press. [PubMed: 22774929]
- Van Lancker D, & Sidtis JJ (1992). The identification of affective-prosodic stimuli by left-and right-hemisphere-damaged Subjects: All errors are not created equal. *Journal of Speech, Language, and Hearing Research*, 35(5), 963–970.
- Wildgruber D, Ackermann H, Kreifelts B, & Ethofer T (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. *Progress in Brain Research*, 156, 249–268. [PubMed: 17015084]
- Wright AE, Davis C, Gomez Y, Posner J, Rorden C, Hillis AE, & Tippett DC (2016). Acute Ischemic Lesions Associated With Impairments in Expression and Recognition of Affective Prosody. *Perspectives of the ASHA Special Interest Groups*, 1(2), 82–95. [PubMed: 28626799]
- Zahn R, Moll J, Iyengar V, Huey ED, Tierney M, Krueger F, & Grafman J (2009). Social conceptual impairments in frontotemporal lobar degeneration with right anterior temporal hypometabolism. *Brain*, 132(3), 604–616. [PubMed: 19153155]
- Zatorre RJ, Belin P, & Penhune VB (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences*, 6(1), 37–46. [PubMed: 11849614]

A. Processes Underlying Affective Prosody Recognition



B. Processes Underlying Affective Prosody Production

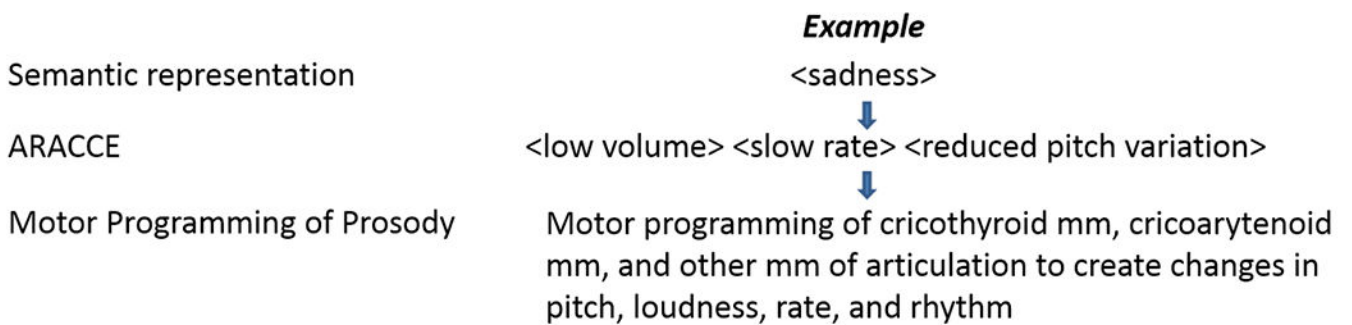


Figure 1.
A proposed cognitive architecture underlying affective prosody recognition (Panel A) and production (Panel B)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

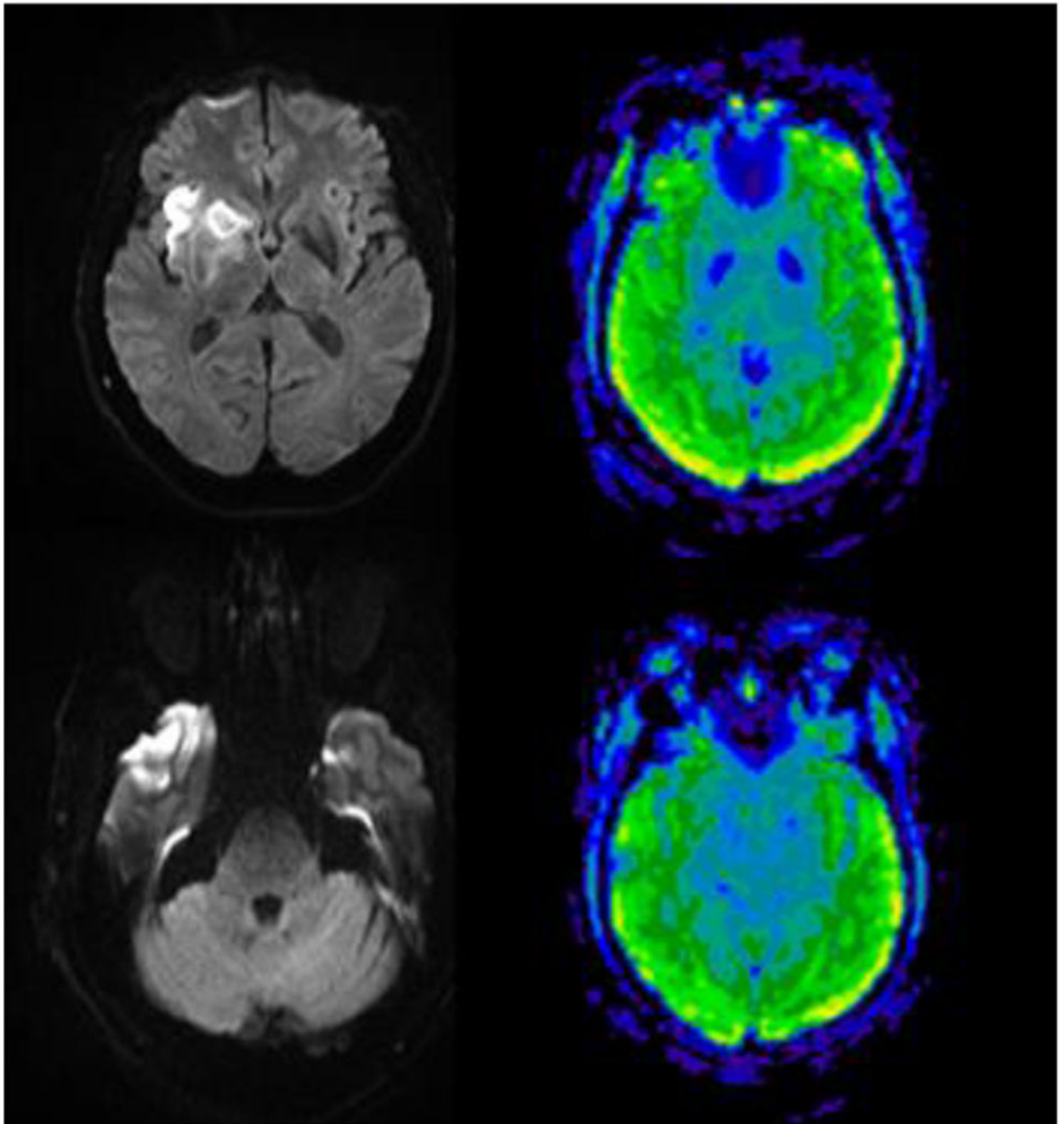


Figure 2. MRI Diffusion-weighted image (left) showing acute infarct in anterior temporal cortex and insula, without hypoperfusion on Arterial Spin Labelling Perfusion Imaging (right) in Case 1, who had impaired extraction of acoustic features from speech. All scans are in radiological convention (right hemisphere on the left).

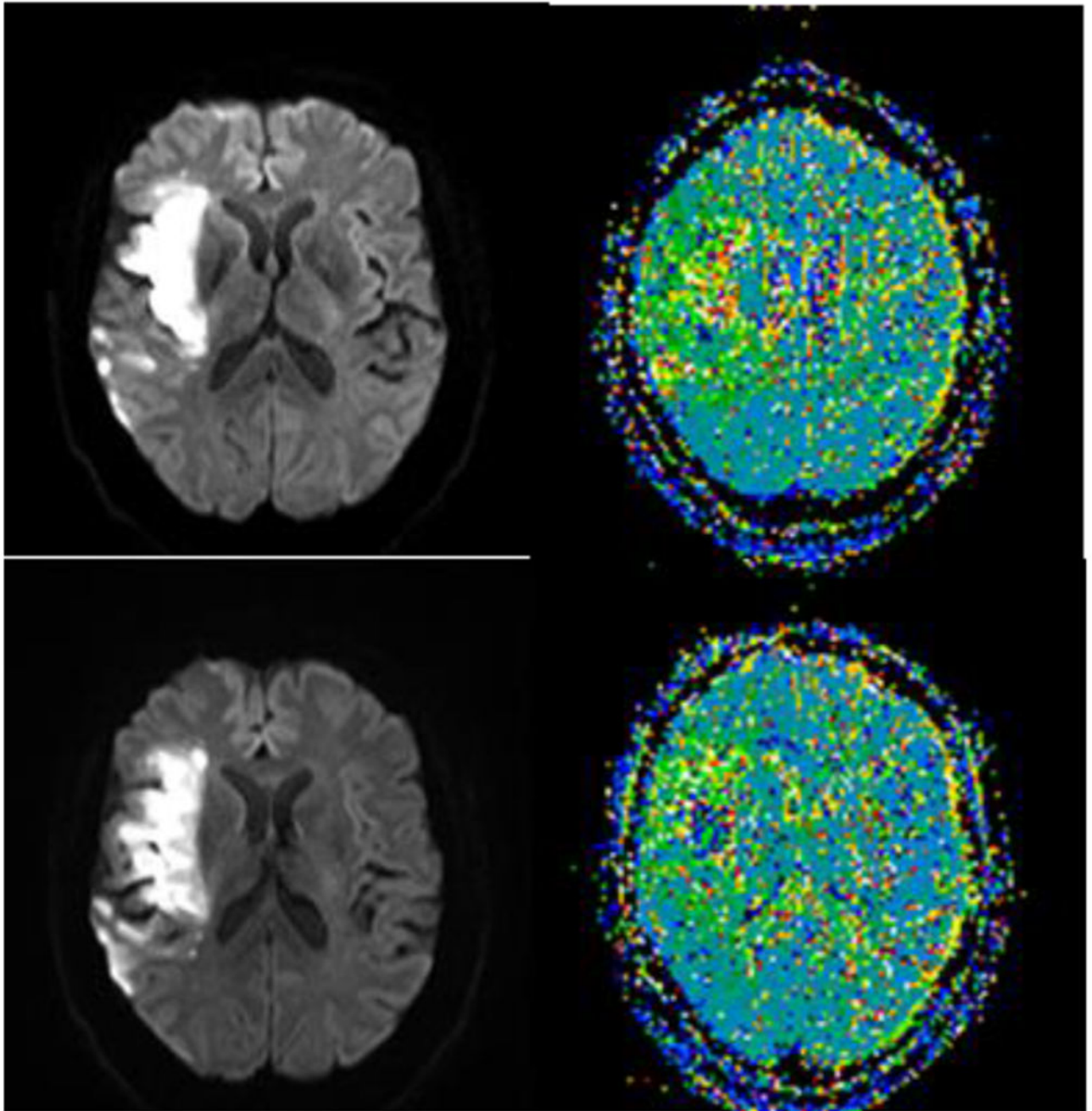


Figure 3. MRI Diffusion-weighted image (left) showing acute infarct in anterior and posterior superior temporal cortex and insula, without significant hypoperfusion beyond the infarct on bolus-tracking Perfusion Weighted Imaging (right) in Case 2, who had impaired extraction of acoustic features from speech.

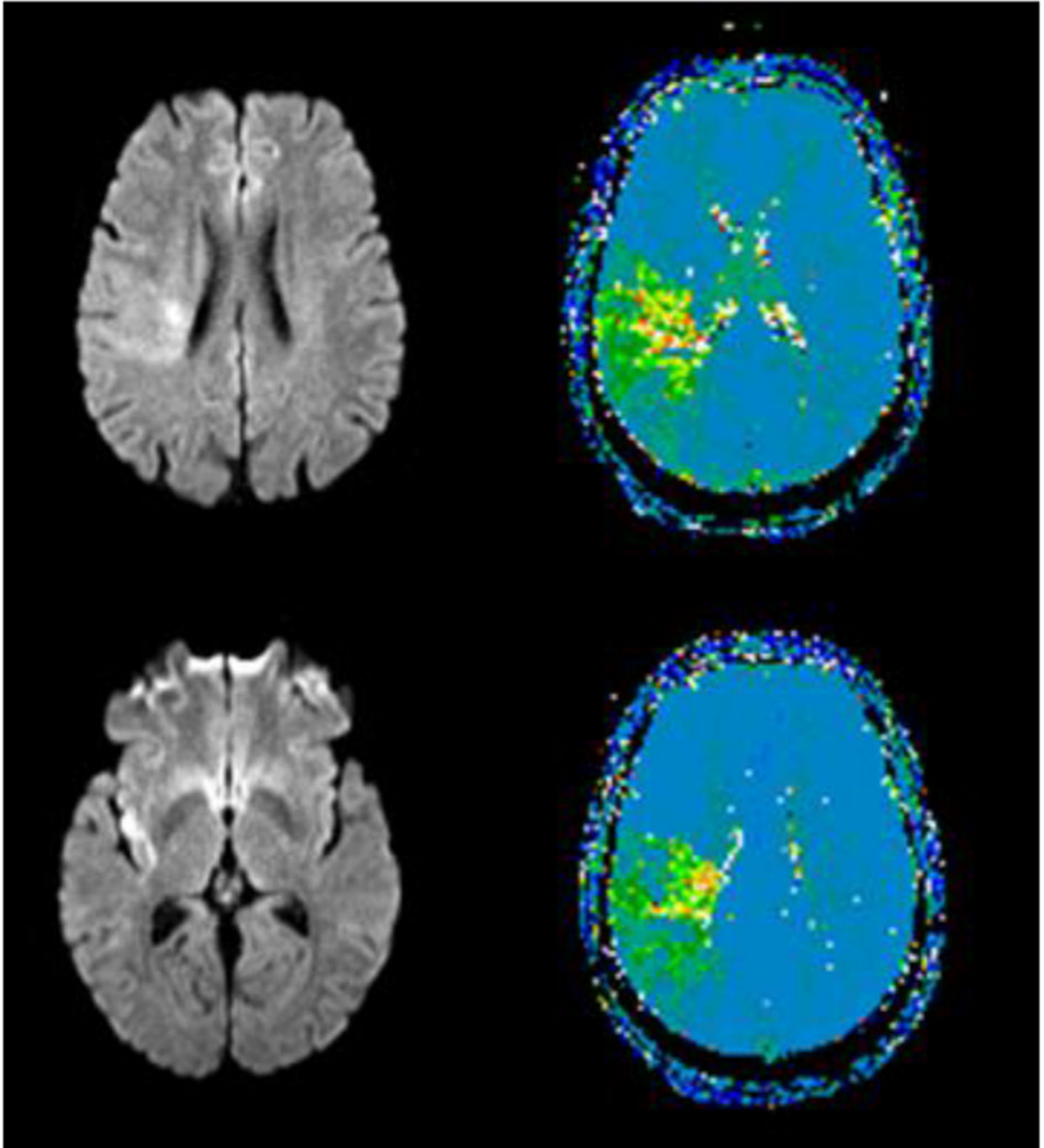


Figure 4. MRI Diffusion-weighted image (left) showing acute infarct in right parietal white matter and insula, with significant hypoperfusion beyond the infarct in parietal cortex on bolus-tracking Perfusion Weighted Imaging (right) in Case 4, who had impaired access to the ARACCE in producing speech. Green areas have significant hypoperfusion (>4 sec delay in time to peak arrival of contrast, relative the homologous area on the left)

Table 1.

Scores on the Prosody Assessments by the Four Cases and the Controls (Given in percent correct for matching and identification tasks and given as coefficient of fundamental frequency (F_0 -CV%) for the production tasks).

Assessment	Case 1	Case 2	Case 3	Case 4	Controls mean \pm SD & number
Identification of acoustic features in sentences	33%	52%	83%	100%	82.4 \pm 4.4 (n= 22)
Identification of features to match emotion words	85%	83%	67%	83%	93.1 \pm 5.4 (n= 7)
Identification of emotions from sentences	50%	29%	30%	85%	90.0 \pm 7.0 (n= 22)
Matching emotions to synonyms	100%	92%	92%	88%	93.6 \pm 5.4 (n= 8)
Matching non-emotional abstract synonyms	100%	92%	88%	100%	98.6 \pm 4.2 (n= 22)
Matching emotions to situations	100%	95%	100%	100%	98.3 \pm 2.6 (n= 6)
Matching emotions to faces	83.9%	DNT	82%	92.5%	89.3 \pm 6.6 (n=40)
Sentence reading without cues (F_0 -CV%)	31.0	24.3	34.6	11.7	24.31 \pm 6.2 (n=10)
Sentence reading with cues (F_0 -CV%)	41.6	32.3	DNT	21.2	19.78 \pm 2.1 (n= 10)
Sentence repetition (F_0 -CV%)	32.3	DNT	29.4	26.5	23.7 \pm 2.6 (n=60)

* DNT= did not test