



## Review

## Subcortical processing in auditory communication

Alessia Pannese<sup>a, b, c, \*</sup>, Didier Grandjean<sup>a, b</sup>, Sascha Frühholz<sup>a, b</sup><sup>a</sup> Swiss Centre for Affective Sciences, University of Geneva, Switzerland<sup>b</sup> Neuroscience of Emotion and Affective Dynamics Laboratory, University of Geneva, Switzerland<sup>c</sup> Département des Neurosciences Fondamentales, Centre Médical Universitaire, University of Geneva, Switzerland

## ARTICLE INFO

## Article history:

Received 3 February 2015

Received in revised form

23 June 2015

Accepted 1 July 2015

Available online 8 July 2015

## Keywords:

Auditory communication

Subcortical processing

Ascending auditory pathway

Affective vocalisation

## ABSTRACT

The voice is a rich source of information, which the human brain has evolved to decode and interpret. Empirical observations have shown that the human auditory system is especially sensitive to the human voice, and that activity within the voice-sensitive regions of the primary and secondary auditory cortex is modulated by the emotional quality of the vocal signal, and may therefore subserve, with frontal regions, the cognitive ability to correctly identify the speaker's affective state. So far, the network involved in the processing of vocal affect has been mainly characterised at the cortical level. However, anatomical and functional evidence suggests that acoustic information relevant to the affective quality of the auditory signal might be processed prior to the auditory cortex.

Here we review the animal and human literature on the main subcortical structures along the auditory pathway, and propose a model whereby the distinction between different types of vocal affect in auditory communication begins at very early stages of auditory processing, and relies on the analysis of individual acoustic features of the sound signal. We further suggest that this early feature-based decoding occurs at a subcortical level along the ascending auditory pathway, and provides a preliminary coarse (but fast) characterisation of the affective quality of the auditory signal before the more refined (but slower) cortical processing is completed.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

Emotional stimuli command vast cognitive resources and benefit from privileged processing across a distributed network of brain regions (Rolls, 1999). Much of the emotional quality of the human voice can be inferred by the physical features of the vocalisation, and by the sonic properties of one's speech. Nonverbal expressions, for example, communicate affective meaning and information about the speaker's state of mind independently of linguistic cues, as shown by cross-cultural studies (Scherer et al., 2001; Pell et al., 2009). The human auditory system is especially sensitive to the human voice. Specific voice-selective areas have been repeatedly reported in the right and left superior temporal

cortex (Belin et al., 2000), with peaks of activation varying considerably across individuals. The networks implicated in the processing of vocal affect includes superior temporal, inferior frontal, and limbic regions (Grandjean et al., 2005; Sander et al., 2005; Schirmer and Kotz, 2006; Bach et al., 2008; Wiethoff et al., 2009; Leitman et al., 2010; Ethofer et al., 2012; Fecteau et al., 2007; Frühholz et al., 2012; Frühholz and Grandjean, 2013a; Wildgruber et al., 2009 for a review). Within the auditory cortex, the voice-sensitive areas have been shown to respond more strongly to emotional as opposed to neutral voices (Frühholz and Grandjean, 2013b), suggesting that information about vocal emotion is already available at the perceptual level of processing. It is however unclear whether this information may also be available prior to reaching the sensory cortex. Hence, a central question in auditory affective communication is at which level of processing the human brain extracts the earliest information relevant for the identification of the voice's affective quality.

Recent models (Schirmer and Kotz, 2006; Wildgruber et al., 2009; Brück et al., 2011) have emphasised the role of the auditory cortex, and proposed that emotionally relevant acoustic information is extracted in the superior temporal gyrus and integrated in

List of abbreviations: BG, basal ganglia; CN, cochlear nucleus; EEG, electroencephalography; fMRI, functional magnetic resonance imaging; IC, inferior colliculus; MGB, medial geniculate body of the thalamus; PAG, periaqueductal grey; SOC, superior olivary complex; STRF, spatio-temporal receptive field

\* Corresponding author. Swiss Centre for Affective Sciences, University of Geneva, Switzerland.

E-mail address: [alessia.pannese@unige.ch](mailto:alessia.pannese@unige.ch) (A. Pannese).

the superior temporal sulcus. Here we elaborate on these models by focussing on the first stage (the extraction), and propose that the earliest emotion-relevant information may also be available prior to the cortical level, at early stages of auditory perceptual processing. For the purpose of this work, we use the term 'perception' to denote a modality-specific sensory processing, independently of whether or not this is reportable and/or consciously perceived – as, for example, in cortical deafness (Graham et al., 1980).

Our model builds on the notion that emotional states can modulate vocal features of the human voice (e.g. intensity of fundamental frequency, energy distribution across frequency bands, distribution of harmonic frequencies) through changes in the peripheral vocal production structures (Banse and Scherer, 1996). Early animal studies have shown a link between the motivational state and the physical structure of the vocalisations emitted in a social context (motivation-structural rules, Morton, 1977). For example, birds and mammals use harsher (covering a wider-frequency band, as opposed to pure-tone-like), lower-frequency vocalisations in hostile contexts compared to friendly settings (Morton, 1977). It has been proposed that these primitive aspects of affective expression have been retained even after the evolutionary transition from basic animal vocalisation to the more sophisticated human language (Grandjean et al., 2006). For example, the physiological changes induced by emotional arousal (e.g. in muscle tone, heart rate, and respiration) may affect the characteristics of the human voice, resulting in an emotional modulation of the vocal expression (Bachorowski, 1999; Bachorowski and Owren, 2003).

Studies have shown that some emotional states share similar influence on vocal features. For example, both in speech (Scherer et al., 1991; Banse and Scherer, 1996; Johnstone and Scherer, 2000) and non-speech vocalisations (Szameitat et al., 2009, 2011; Sauter et al., 2010), anger, fear, and joy have all been associated with increased pitch. However, despite these commonalities, different emotional states are typically associated with distinct overall patterns (Patel et al., 2011), that is, distinct combinations of features. For example, anger has been associated with increased pitch and loudness; fear with increased pitch and speech rate; joy with increased pitch, loudness, and speech rate; sadness with decreased pitch, speech rate, and loudness (Scherer et al., 1991; Banse and Scherer, 1996; Johnstone and Scherer, 2000; Szameitat et al., 2009, 2011; Sauter et al., 2010). The resulting unique vocal profile helps to differentiate the affective quality of the vocalisation, and respond with the appropriate behaviour (e.g. differentiating anger from fear enables responding appropriately, i.e. approaching rather than avoiding, Marsh et al., 2005).

At the most basic level, affective vocalisations consist of sounds, hence their perception relies on the ascending (i.e. afferent) pathway of the auditory system. This pathway originates from the auditory nerve, and terminates in the auditory cortex, via several intermediate relay points (see Gelfand, 2009 for overview). Specifically, auditory nerve fibres terminate in the cochlear nucleus, where they synapse with neurons projecting (mainly contralaterally) to the superior olivary complex. Neurons originating from the olivary complex ascend via the lateral lemniscus to reach the inferior colliculus, and the medial geniculate body of the thalamus, which is the last subcortical station of the auditory pathway. From the thalamus, neurons target the auditory cortex, located in the transverse temporal gyrus. Various structures within the ascending auditory pathway have been shown to respond to the physical features of acoustic signals. Starting from the auditory nerve, virtually all the nuclei along the subcortical auditory system are sensitive to specific sound properties. Here we will discuss specifically the cochlear nucleus (CN), the inferior colliculus (IC), and the medial geniculate body (MGB) of the thalamus. Additional

structures outside the ascending auditory pathway will also be discussed in virtue of their tight anatomical and/or functional connection with the auditory perceptual system or known involvement with affective communication. These are the periaqueductal grey (PAG), the basal ganglia (BG), and the amygdala.

### 1.1. Cochlear nucleus

The cochlear nucleus (CN) is a highly conserved structure found in all vertebrates (Baird, 1974). It is comprised of diverse and highly specialised neuronal populations (Ramón y Cajal c.1900; Lorente de Nó, 1933, 1981) differing in morphology (Harrison and Warr, 1962; Osen, 1969a; and 1969b; Brawer et al., 1974; Heiman-Patterson and Strominger, 1985; Adams, 1986), response properties (Pfeiffer and Kiang, 1965; Pfeiffer, 1966; Evans and Nelson, 1973; Shofner and Young, 1985; Blackburn and Sachs, 1989; Winter and Palmer, 1990; Marsh et al., 2006), and connectivity with the auditory nerve (Lieberman, 1991, 1993). Each neuronal population in the CN exhibits distinct projections (summarised in Cant and Benson, 2003), some ipsilateral (Cant and Casseday, 1986), others contralateral (Warr, 1966, 1969; Winer and Schreiner, 2005), reflecting the distinct functions of the CN anteroventral (involved in sound localisation) and posteroventral regions (involved in sound discrimination) (Eggermont, 2001).

At the functional level, evidence for a role of the CN in auditory vocal communication comes from animal studies showing that neurons in the rats' ventral CN code the syllables' voice-onset time (Clarey et al., 2004). These findings corroborate earlier results from intracellular recordings (Stevens and Wickesberg, 1999; Pressnitzer et al., 2000) and are consistent with the hypothesis that the hyperpolarisation observed in the CN is due to the stimulation of inhibitory sidebands, which reduces the probability of action potential generation and enhances the salience of the voice onset (Paolini et al., 2001). Although hypothetical, the notion that the salience of vocal features may be modulated by activity in the CN – hence, very early in the afferent auditory pathway – is compatible with anatomical evidence of direct projections from this nucleus to the medial geniculate body (MGB) of the thalamus, i.e. bypassing the inferior colliculus (Woollard and Harpman, 1940; Strominger and Strominger, 1971; Strominger, 1973; Strominger et al., 1977; Malmierca et al., 2002; Anderson et al., 2006). Future studies might investigate the functional role of these pathways in processing voices and voice sound features. Besides anatomical evidence from anterograde tracers, this latter study also showed functional evidence for direct projection: neurons in the medial MGB respond to acoustic stimuli with a latency of less than 6.5 ms, which is too short for a transit through the inferior colliculus (IC), and only possible in the presence of a direct (high-speed) pathway from the CN to the MGB (bypassing the IC). This pathway has been further characterised as constituted by (mostly) contralateral fibres originating predominantly from multipolar T-stellate cells of the ventral (and fewer cells of the dorsal) division of the CN, and terminating on the medial MGB, with most cells sending collateral axons (hence, presumably the same information) to the IC as well (Anderson et al., 2006).

A possible implication of a direct projection that bypasses the multiple auditory nuclei (e.g. superior olive, lateral lemniscus, IC) and enables early (short-latency) medial MGB exposure to the acoustic information is that it provides the thalamus with a "preparatory signal" that facilitates further processing of the acoustic stimulus, whose information will eventually be carried more fully through the main (i.e. passing through IC and all the auditory nuclei) ascending auditory pathway (Anderson and Linden, 2011). Since T-stellate cells display a narrow frequency tuning and encode stimulus envelope (Oertel et al., 2011), it has been suggested that

Download English Version:

<https://daneshyari.com/en/article/6287263>

Download Persian Version:

<https://daneshyari.com/article/6287263>

[Daneshyari.com](https://daneshyari.com)