

Research Articles

Voice-Sensitive Areas in the Brain

A Single Participant Study Coupled With Brief Evolutionary Psychological Considerations

George Varvatsoulas*^a

[a] Newham University Centre, London, United Kingdom.

Abstract

This empirical single-participant fMRI case study partially replicates Belin, Zatorre, Lafaille, Ahad, and Pike (2000) research on voice-selective areas in the auditory cortex. It hypothesises that brain areas, sensitive to human vocal sounds, show greater neural activation than non-human ones. A 1X3 ANOVA design was used in this study contrasted by two conditions: sound vs. silence and voice vs. non-voice. The findings supported the hypothesis, noting also possible individual differences to the degree of voice activation in both hemispheres. Suggestions for a future replication could discuss voice/non-voice and speech/non-speech neuronal activation in the brain, auditory and visual neural responsiveness to voice and face modalities, and evolutionary assumptions in regard to sound- and voice-selective reactivity.

Keywords: superior temporal sulci, superior temporal gyri, evolutionary psychology

Psychological Thought, 2014, Vol. 7(1), 66–79, doi:10.5964/psyc.t.v7i1.98

Received: 2013-10-29. Accepted: 2013-12-27. Published (VoR): 2014-04-30.

Handling Editor: Marius Drugas, University of Oradea, Romania

*Corresponding author at: Newham University Centre, Stratford Campus, Welfare Road, London, E15 4HT, UK. E-mail: George.Varvatsoulas@newham.ac.uk



This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Human communication depends on voice/non-voice and speech/non-speech sounds associated with cortically-specific areas in the brain (Binder, Frost, Hammeke, Bellgowan, et al., 2000). Through voice, humans can analyse information about an individual's sex, age, emotions, even if one's language is not familiar to the language the person is listening to (Ward, 2010). Human voice is a powerful tool for speech processing in terms of identification of voice-specific vs. non-voice specific sounds. Human voice is activated via brain regions which are voice-selective and can be found in the auditory cortex (Belin, Fecteau, & Bédard, 2004; Belin, Zatorre, & Ahad, 2002; Belin, Zatorre, Lafaille, Ahad, & Pike, 2000). Recognition of an individual is also a voice-related characteristic, though one's identity is more easily accessed by looking at one's face, instead of listening to one's voice (Hanley, Smith, & Hadfield, 1998).

Voice-selective areas are auditory cortical regions triggered not only by human sounds, but by socially-related information, linked to other sensory systems as well, such as vision, touch and proprioception (Tsao, Freiwald, Tootell, & Livingstone, 2006). Voice-selective areas are located in the superior temporal lobes of the auditory cortex – function of the 'what' sourceⁱ –, and are activatedⁱⁱ during the operation of human vocal sounds (Vuilleu-

mier, 2005) where responsiveness to vocal sounds exhibits greater neural sensitivity (Belin et al., 2000)ⁱⁱⁱ. By voice-selectivity it is meant that the brain reacts more accurately in the hearing of human sounds compared to non-human ones (Belin et al., 2000).

Human voice differs from non-human in that it refers to vocalised speech (Levy, Granot, & Bentin, 2001). Superior temporal regions respond to voice-selectivity in regard to vocalised-specific information as well as affective communication^{iv} (Ethofer, Anders, Erb, et al., 2006; Ethofer, Anders, Wiethoff, et al., 2006). Voice-selective areas are found in the same sensory tracts both in humans and primates (Petkov et al., 2008).

Voice-selective areas are activated in voice-selective regions sensitive to the increase and/or decrease of human vocal sounds when modulated by attention-related factors, such as spatial and selective attention to tasks (Hashimoto, Homae, Nakajima, Miyashita, & Sakai, 2000). That means superior temporal lobes react to human sounds by demonstrating higher receptivity compared to non-human ones. Vocal-sound areas can be speech and/or non-speech, for the neural activity they exhibit is higher compared to non-vocal areas (Belin et al., 2000). That is explained by the capacity of speech/non-speech sensory areas in the brain, whereby communication between conspecifics is emergent in both verbal (left superior temporal gyrus and bilateral posterior parietal) and non-verbal regions (right dorsolateral prefrontal cortex – DLPFC – and bilateral superior frontal gyrus) of the brain (Rothmayr et al., 2007; Shultz, Vouloumanos, & Pelphrey, 2012).

Acoustic features of both vocal (speech) and non-vocal (non-speech) stimuli are not processed the same way from the human brain (Joanisse & Gati, 2003). Voice (speech/non-speech)-selective areas are also located along the superior temporal sulci (STS) region of the right auditory cortex, indicating that neural projections in the STS are involved in the analysis of complex acoustic/non-acoustic features coming from vocal/non-vocal stimuli (Belin et al., 2000; Binder, Frost, Hammeke, Rao, & Cox, 1996). This is something supported by lesion studies, with patients who though could not recognise human voice, were able to identify the expressed content of a voice (Benzagmout, Chaoui, & Duffau, 2008). Non-speech, non-vocal and non-acoustic features of sound refer also to auditory and non-auditory stimuli entering the same areas found in temporal lobes – as explained above – both for those with sound – or hard hearing, as well as for those with non-hearing capacity at all (Traxler, 2012).

Findings that support STS bilateral activation for human vocal sounds show that such voice-selectivity is more evident when context-specific voices are heard, whereas the opposite is true for non-speech, non-vocal and non-acoustic features of sound (Patterson & Johnsrude, 2008). Speech processing refers to context-specificity and can be observed across several neural loci in the STS (Uppenkamp, Johnsrude, Patterson, Norris, & Marslen-Wilson, 2006), though little is known about its co-functional character to voice processing areas, when non-vocal stimuli are involved, such as pitch discrimination and pitch matching abilities in terms of amplitude-modulated stimuli of white noise^v (Binder et al., 1999; Moore, Estis, Gordon-Hickey, & Watts, 2008; Steinschneider, 2012; Zatorre, Evans, Meyer, & Gjedde, 1992). An explanation could be that, speech processing relates to the production of articulated sounds, which can be understood in terms of linguistic and voice content (Bonte, Valente, & Formisano, 2009). Linguistic content is mainly located in the left (Morillon et al., 2010), whilst voice-content in the right hemisphere (Grossmann, Oberecker, Koch, & Friederici, 2010).

Human vocal sounds correspond to voice-selective areas in the auditory cortex, however, what about auditory stimuli that look similar to human voice, such as non-sense sounds or sounds associated to verbal communication, i.e., incomplete words or letters missing from words? Would they demonstrate the same neural activation? According to relevant research (Bélizaire, Fillion-Bilodeau, Chartrand, Bertrand-Gauvin, & Belin, 2007), the answer is that

natural human sounds elicit cortical reactions at the STS, whereas voice-like stimuli reduce such reactions. The understanding of this finding is that voice-selective regions are contrasted to stimuli of human voice resemblance.

The literature on voice-selectivity apart from exploring human voice areas, explores also possible multimodal integration between auditory and visual systems (Campanella & Belin, 2007). Such integration, though it takes into account that auditory and visual cues are processed differently, it brings together voice and face selectivity associated with the speaker's identity (Ward, 2010). That does not mean that auditory and visual parts are cortically related, but that the sensory information processed can simultaneously activate both cortices, if the person is familiar to the listener (Ethofer, Anders, Wiethoff, et al., 2006). Voice-selective regions in the auditory cortex might be seen as parallel to face-selective areas in the visual cortex, such as remembering one's voice by seeing one's face or remembering one's name by the way one is walking or sitting (Kanwisher, McDermott, & Chun, 1997). In this way, auditory and visual sensory systems can be functionally understood as being commonly designed for, so that to both react on human vocalisations and face recognition stimuli (Sweller, 2004).

Through functional magnetic resonance imaging (fMRI), cognitive neuroscience can identify brain regions related to the recruitment of voice and non-voice processing of sound (Vouloumanos, Kiehl, Werker, & Liddle, 2001). Voice-sensitive areas in the brain provide the cortical tools for the comprehension of auditory information (von Kriegstein, Eger, Kleinschmidt, & Giraud, 2003). Voice perception plays a significant role in identifying an individual speaking and the emotions one's voice is carrying. Not much is known about its neural precursors (Belin et al., 2004; Imaizumi et al., 1997). By obtaining a fuller neural picture of the voice areas, human auditory cortex can be better understood in terms of connectivity with other brain parts (Belin et al., 2000).

The present report will be a partial replication of the Belin et al.'s (2000) study. Theoretically, this study, follows a study by Varvatsoulias (2013) which presents and discusses the neuroscientific background on fMRI and its importance in the scanning of neuronal activation. This study will investigate voice-selective regions in the human auditory cortex that mainly react to human vocal sounds. Belin et al.'s (2000) study is a seminal paper, for they were the first to point out that:

1. Superior temporal sulci areas in the right hemisphere, such as the anterior section of the temporal part – crucial for speaker's identity – and, the central section of the anterior extension of Heschl's gyrus (HG), and the posterior section of Heschl's gyrus^{vi}, show greater reactivity to human vocal sounds.
2. STS voice-selective response does not necessarily depend on speech vocal stimuli
3. Voice-selective areas can also respond to sounds of non-human origin
4. Voice-sensitive areas can be selectively recruited by a combination of high and low voice-featured frequency
5. By voice-responsiveness is not meant that voices are specifically elicited by voice-selective areas in the brain
6. Decreased neural activity in the auditory cortex can be a result of participants' behavioural changes during performance on voice-perception tasks
7. Voice-selective regions may be regarded as analogous to face recognition processing

Rationale

Voice-selective areas are differently activated from non-voice. The neural activation differs also among those areas in that it accounts for specialised and non-specialised processing of auditory information (Whalen et al., 2006).

Aim and Hypothesis

The aim of this report is to study the neural underpinnings of voice-selective perception by hypothesising that auditory cortex areas are specifically sensitive to human vocal sounds than to other sounds^{vii}.

Method

Participants

A healthy 24 years old male participant, native Greek and fluent English speaker^{viii}.

Design

1X3 repeated measures, event-related fMRI, ANOVA design; IVs: silence, voice, non-voice; DV: brain activity. Two conditions will be contrasted: sound (voice + non-voice) vs. silence and voice vs. non-voice.

Stimuli

Auditory stimuli of sounds were presented at a sound-pressure level of 88-90dB. Stimuli were composed of sounds from a variety of sources in sixteen blocks of similar overall energy (RMS) by the use of Mitsyn (WLH) and CoolEdit Pro. Sounds were played for both 'voice' and 'non-voice' category with a rest period in-between. Sounds for the 'voice' category were human vocalisations, not essentially of speech, such as coughs, random utterances, nonsense words, or singing, whereas for the 'non-voice' were environmental sounds from nature, animals, or technology.

Procedure

The participant was asked to close his eyes and listen to the presentation of various sounds. In both 'voice' and 'non-voice' categories sixteen blocks of sounds were delivered at onset times of '15 45 105 135 165 195 255 285' and '30, 60, 90, 120, 180, 210, 240, 270' seconds respectively. In the 'voice' category the subject passively listened to vocal sounds – speech or non-speech – coming from talkers of different age and gender; in the 'non-voice' category energy-matched non-vocal sounds were presented to the participant.

fMRI Recording Parameters

A 3-T MRI Siemens trio scanner with a head coil of eight-channel array was employed. Standard gradient echo provided functional images of the whole brain via an echoplanar sequence of TR 3000ms, TE 33ms, 39 slices, 3x3x3 mm voxel size, 64x64 matrix, and a 90 degrees flip angle. Every experimental sequence was axially oriented, so pictures to be obtained successively. A 3-D whole brain T1-weighted anatomical mapping was obtained in a 1x1x1mm high resolution with a sequence of TR 1900ms, TE 5.57ms, and 11degrees flip angle.

Pre-Processing

Re-alignment of functional images was carried out, so participant's motion effects to be eliminated. Re-alignment was normalised according to the Montreal Neurological Institute (MNI) standard space, whereas it was also smoothed by the use of a 4MM Gaussain kernel. The image of the participant's brain was normalised with a T1-weighted high-resolution, according to the identical MNI space, so results to be efficiently presented.

Statistical Methods for Data Analysis

The general linear model (GLM) was employed for the analysis of data. Regressors' inputs were generated to the model by following the convolution of the on-off block cycle. This corresponded to the Voice/Non-Voice presentation

period of stimuli with reference to hemodynamic function, so dispersion and delay of the events following stimuli to be measured in regard to neural responses. Contrasts been used were the 'Voice + Non-Voice overall effect vs. Silence' (F Contrast), and regions responding to the stimulation of 'Voice over Non-Voice' (t Contrast). Significant activation of voxels at a $p < .05$ statistical threshold was assumed, being family-wise error (FWE) corrected for, again at a $p < .05$ threshold.

Results

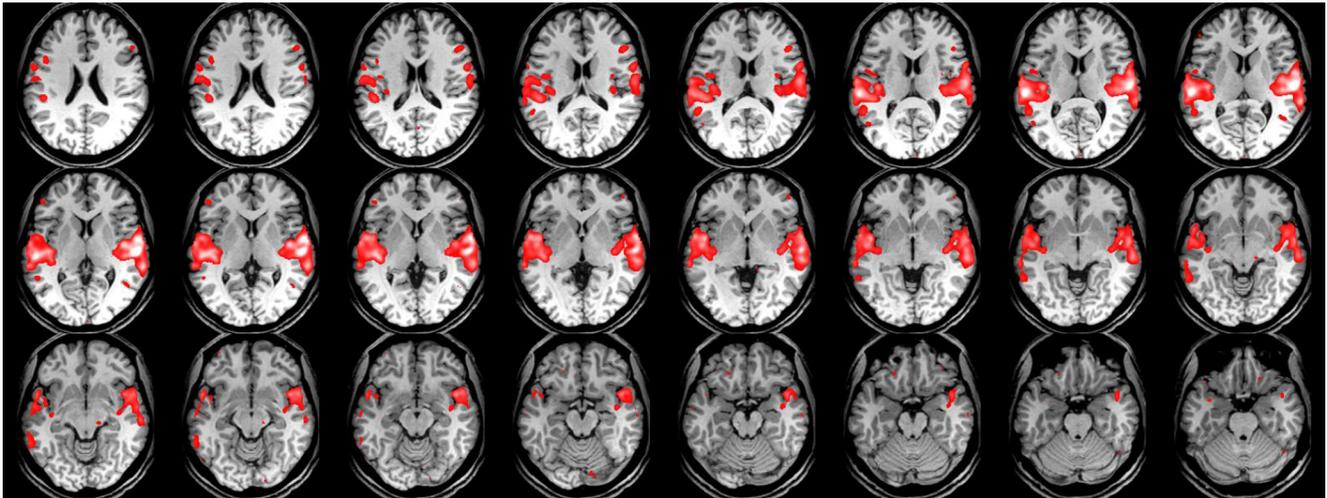


Figure 1. Images correspond to brain areas scanned by fMRI. Sound areas in the superior temporal gyrus area are slightly more demonstrated in the right cerebrum compared to silence ones which are mainly located in the superior temporal gyrus area of the left cerebrum.

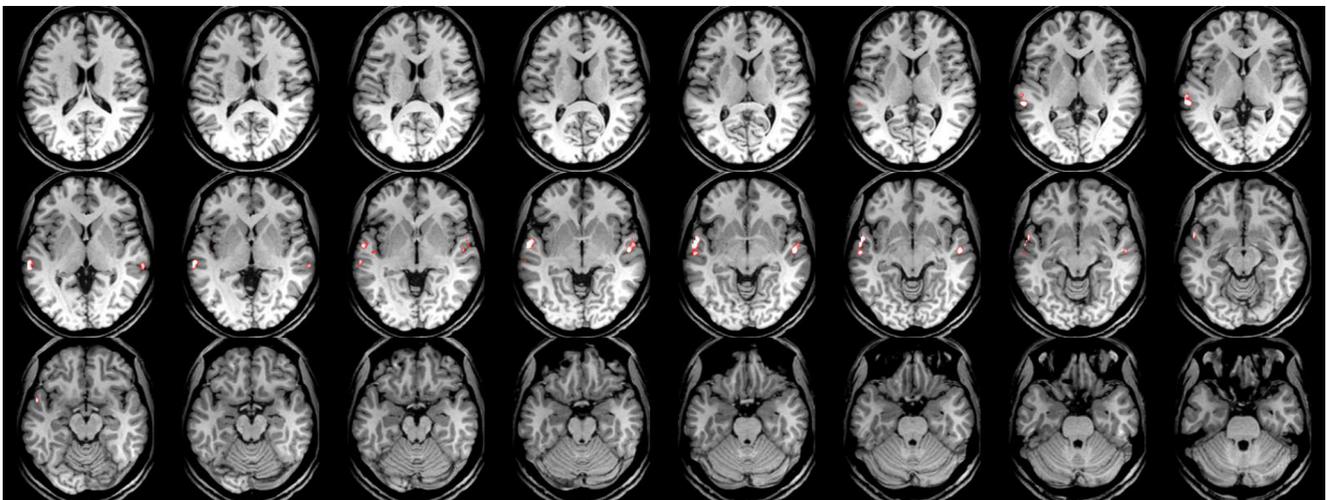


Figure 2. Images correspond to brain areas scanned by fMRI. Voice neural response in the superior temporal sulcus area is clearly located in the left cerebrum, whereas non-voice one is located in the superior temporal sulcus area of the right cerebrum.

The critical value $F = 16.4$ ($p < .05$) shows significantly higher activation, bilaterally in the superior temporal gyri (STG, with $x = 58$, $y = -18$, and $z = 12$ coordinates), in the sound (voice + non-voice) category compared to silence.

The critical t-value ($t > 5.16$, $p < .05$) shows significantly greater activation, bilaterally in the STS (with $x = 76$, $y = 50$, and $z = 16$ coordinates), for human vocal sounds compared to non-human (see [Figure 1](#) and [Figure 2](#)).

Discussion

Having explained the reasons in footnote viii why there has been conducted a single participant study, the hypothesis is supported in the light of the results obtained. Auditory cortex is specifically activated by human vocal sounds compared to non-human. Voice-sensitive areas in the brain – speech or non-speech – are bilaterally activated in the STS, whereas sound regions bilaterally in the STG. Auditory areas in the brain demonstrate greater sound selectivity than voice/non-voice.

This partial replication agrees to the findings argued by the paper ([Belin et al., 2000](#)) that the superior temporal sulci are mainly voice- and not non-voice selective. Findings of this fMRI study portray areas in the auditory cortex with higher neural activation to sound compared to voice-sensitive areas. Findings show that the brain consists of specific auditory regions which are voice-selective. Findings exhibit that voice-selectivity is not only sensitive to sounds of human origin. Findings conclude that voice-selective areas respond to human and not non-human voice.

According to the findings outlined, human vocalisations activate the left hemisphere compared to alternate sound categories. This confirms that voice sounds elicit neural arrays mainly of human-voice responsiveness ([Charest et al., 2009](#)). In listening to sounds, there is more cortical tissue activated in the right hemisphere than in the left, as well as better neural response compared to silence. Neural responsiveness to sounds can also be subject to the intensity of pitch during which sounds are heard ([Palmer & Summerfield, 2002](#)).

In this single participant study, voice-sensitive areas in the STS were shown to be mainly located in the left hemisphere, and not in the right, as the paper ([Belin et al., 2000](#)) has argued. That, though does not render the paper's ([Belin et al., 2000](#)) findings ineffective, for individual differences can also be observed, it does however offer the hypothesis of little laterality for vocal stimuli in the brain and in particular to the auditory areas the fMRI scan has shown that during the process they become activated, i.e. in the STS area in the left cerebrum. Such hypothesis may offer continuation to the debate between hemispheric asymmetry in vocal processing and speech or linguistic processing, with or without the involvement of non-verbal affect bursts^{ix}, during the accomplishment of auditory information ([Grossmann et al., 2010](#); [Scherer, 1995](#)). In line with what has been argued, as to the hypothesis stated above, mismatch negativity measurements in studies of event-related potentials (ERPs) seem to be in favour of such hypothesis. That is to say that in auditory perception, what is of crucial role is the presence or not of attention so that the linguistic experience, when obtained, to determine whether responses heard relate to speech or non-speech sounds ([Näätänen, Paavilainen, Rinne, & Alho, 2007](#)). By that it is meant that what is retained in the auditory sensory memory, and therefore forms part of auditory perception, is the cognitive comprehension of frequent – standard – and infrequent – deviant – sounds in terms of mismatch negativity (MMN) measurements so that auditory areas to be further studied ([Näätänen, 1992](#))^x. Moreover, and in reference to the above hypothesis, the rationale behind MMN is that such studies provide a clearer understanding about the ability of the brain to discern stimuli instead behaviours, with the former to be the outcome of factors related to attention issues and/or motivation, and the latter, the effect of actions in the here and now ([Bishop & Hardiman, 2010](#)).

On the other hand, the fact that the study was conducted with only one participant, cannot prove justice to a representative sample. If more participants took part in the study, the degree of same areas activation could vary, suggesting individual differences in the neural response to auditory information (Buchweitz, Mason, Tomitch, & Just, 2009). The participant was male, implying that gender differences in auditory information processing could also be suggested as an important element of discussion to neural activation, if there were to be recruited female participants as well (Brown, 1999).

Also, the participant was not a native English speaker, meaning that if subjects were both native and foreign speakers of English, neural performance could be studied in terms of bilingual proficiency as to the vocal stimuli presented (Fabbro, 2001). Participants, for instance, could read a text in English and then translate and read it in their own language in order for auditory cortex correspondence to be detected (Price, Green, & von Studnitz, 1999). Allocation also of an appropriate number of participants, males and females, native and non-native speakers, will generate more accurate findings and will eliminate confounding variables relevant to the limitations discussed. In such a study, there could also be investigated whether native vs. non-native language proficiency activates same or different cortical regions. In trying a study with more participants, what could further be tested would be auditory triggers in the form of verbal affect bursts associated with human voices, presented in one's native language as well as not. The purpose for such a study could be the research of affective elements of cognition in terms of emotional reasoning, and which parts of the brain are activated, when stimuli are heard in different languages (Belin et al., 2008). Such study could also include high and low-frequency words so that the effect of commonly/not so commonly words in neural responsiveness to be detected. Word frequency has been found to be associated with left inferior frontal activation (Ghee, Soon, & Hwee, 2005). In such a way, high and low word frequency would probably represent different neural arrays in terms of exposure consistency vs. inconsistency (Balota & Chumbley, 1985).

In view to such a study, therefore, what could also be researched as well would be whether the auditory cortex precipitates, or is precipitated by, the neural activation of visual cortex, if participants were *to see the words* when these were uttered to them. A parallel study to that could be looking at familiar faces on a computer and then verbally recalling their names, and vice versa. What could be measured would be the time needed to recall a name versus face, and the neural activation of brain regions during both conditions. In such a study, findings could probably refer to greater or lower neuronal response between voice and face-selective areas. Also, as to such a study, faces of high versus low familiarity could probably extend neural activation mapping to other lobes of the brain too, such as the occipital lobes, whereby seeing a face could suggest increased or decreased activation in neural arrays; something that could also be the case as for words of high and low frequency (Hintzman, 1988). In this way, the study of known vs. less known faces and high vs. low-frequency words could therefore question the effect of commonly used words vs. *not*, in neural receptivity. Word frequency has been found to be associated with left inferior frontal activation (Ghee et al., 2005). In such a way, words of high and low frequency may probably represent different neural arrays when examined in respect to exposure consistency vs. inconsistency (Balota & Chumbley, 1985).

It has also been argued that neural recruitment of information depends on all kinds of sounds whether voice/non-voice, or speech/non-speech (Patterson & Johnsrude, 2008). To such an extent, should a study be conducted could investigate species-specificity areas in the brain by looking at voice/non-voice and speech/non-speech neural responsiveness. What could be examined in that sense would be issues of neural activation between spoken and non-spoken information (Jaramillo et al., 2001), as well as voice and speech-related sounds associated

with literal and abstract linguistic content. What could be explored would be the corresponding cortical activation of species-specificity areas to literal and abstract linguistic content for both voice and speech vocalisations of human origin (Fecteau, Armony, Joannette, & Belin, 2004).

Furthermore, the investigation whether the auditory cortex elicits or elicited by visual cortex activation could be another important chapter researching co-functionality issues between sensory regions (Wheeler, Petersen, & Buckner, 2000). Participants could look at familiar faces, and then asked to recall their names. What could be measured would be the elapsed time from the onset of each face to the recollection of name, and the neural activation of brain regions during both conditions. Findings-to-be may predict greater or lower neural correspondence between voice and face-selective modalities (Campbell, 2008).

Last, but not least, the neural activation to voice-selectivity has been an issue of discussion in evolutionary psychology as well. This is because humans are capable in distinguishing environmental sounds differently to voices of conspecifics. We have observed in this study that areas of sound are more distinctly illuminated compared to areas of human vocal sounds. That may be a reason of evolutionary origin, since human life depends on surrounding sounds, which are many, compared to sounds of human origin, which are relatively less. The latter refers to context-specific adaptive problems, whereas the former to universal acoustic features affecting everyone (Westman & Walters, 1981).

Humans mainly communicate with each other by the use of species-specific vocalisations, whereas communication with the environment involves more complex interplay of sounds (Gygi, Kidd, & Watson, 2007). Voice-selective areas in the brain can be evolutionarily explained in terms of sociality-adaptive problems that had to be resolved in order for ancestral communities to be able to survive (Locke, 2008). Human voice was a costly signal, and a prominent indicator of communication and cooperation in ancestral environments (Fiske & Ofshe, 1970). Adaptive problems, such as group formation, in-group alliances, out-group conflict, competition strategies, mating preferences, social exchange practices, and cheater detection, were an everyday struggle between conspecifics that had to be regulated (Barrett, Dunbar, & Lycett, 2002).

Day-to-day adaptations and survival needs in given milieus were however costlier than those of personal and social focus (Corning, 2000). Adaptive challenges elicited greater importance to survival needs long before humans established interrelationships (Cornwell et al., 2007), something that is argued to be the reason for sound-sensitive areas in the brain (Hester, 2005). In light therefore to evolutionary findings, should a hypothesis be suggested could question adaptations in view to societal needs sound-selective regions have been selected for to demonstrate higher neuronal activation compared to voice-selective regions.

Conclusions

This fMRI report was a single-participant study due to time restrictions to collect data as to the availability of the scanner, as well as that the participant was bilingual, compared to the study that was partially replicated (Belin et al., 2000), which included only an English-speaking sample. Though from a single-participant study cannot be justified findings referring to a particular population of a sample, this study did actually support the hypothesis that human voice elicits greater activation than non-human voice. The present study further supports the hypothesis for a continuation of the debate between hemispheric asymmetry in vocal processing, and speech or linguistic processing, with or without the involvement of non-verbal affect bursts during the accomplishment of auditory information. In that, little laterality for vocal stimuli in the brain could increase the discussion towards understanding

the issue of hemispheric asymmetry as an ongoing neural process for the accommodation of auditory information. Human vocal sounds are more significantly activated compared to non-human, which implies that human voice is context-specific, whereas non-human voice is not. Sound areas in the brain are also more considerably activated compared to silence, in that they refer to a large amount of vocal vs. non-vocal sounds, in that they can be triggered by a richer array of environmental stimuli.

Also, voice-selective areas were found to neural loci bilateral to superior temporal sulci, whilst sound-selective areas were bilaterally activated to superior temporal gyri. In a future replication of this study more participants from both genders should take part; issues of bilingual significance as to neural activation and performance could also be included in the discussion; whereas also, aspects of neural activation and species-specificity of voice/non-voice and speech/non-speech could be looked as well. Last, but not least a further discussion in regards to the evolutionary importance of neuronal activation of sound-receptive regions compared to voice could be investigated, so that adaptive explanations to be elaborated in the discussion.

Notes

- i) *'What' source*: The origin of stimulus-i and its/their relation to the conveyed information.
- ii) Sensory system processes refer to how physical phenomena are perceived. Physical phenomena are stimuli with particular intensity, location and duration. In the emergence of a stimulus, temporal and occipital brain parts are activated to receive the information entering their lobes. Whether auditory or visual information enters those lobes, their anticipation on the perceived stimulus-i depends on the different ways neurons are firing so that the entered sensory information to be interpreted through the brain's cognitive abilities. The triangle between physical phenomena/stimuli, temporal/occipital lobes and cognitive abilities activation through neuronal functions, explains what sensory processes are and how they operate so that external information to be taken in (Kolb & Whishaw, 2008; Krantz, 2012).
- iii) Superior temporal lobes are voice-receptive areas in the brain where acoustic information is intercepted and 'translated'. By 'translated' I mean that auditory information entering the brain can either be understood by superior temporal lobes as emotion-laden (right lobe) or interpretation (the meaning of a particular voice: concepts and ideas) and comprehension (who is talking to whom, who is the person talking to oneself when voice is not followed by vision) (left superior temporal lobe).
- iv) 'Affective information' explains the emotional reaction following the entrance of stimuli in the various sensory tracts of human brain.
- v) White noise is the flat acoustic or non-acoustic tone of signalling – any kind of tone: could be a 'thud', high/low-pitch frequency of a sound or non-sound, etc. – that can either be combined with visual and/or auditory stimuli.
- vi) Heschl gyrus is the auditory cortical centre for hearing. Heschl's gyrus is considered as the morphological feature of the auditory cortex activated in the learning of music: it is thought that the more this auditory part is activated the more one obtains 'musical ear' (Schneider, Scheng, Dosch, Specht, & Rupp, 2002).
- vii) Though, this hypothesis is a general statement for it refers to a single-case study, it has been worded that way for the reason the present research is a partial replication of the study by Belin et al. (2000), which is mentioned above.
- viii) The reason this study has been a single-participant one was because other researchers were also using the same fMRI scanner, and due to time-restrictions –for it was also employed by medical doctors– we had to collect our data quickly. The time spent to collect data from one participant took about two hours, whereas if participants would be more I would probably need days to complete it –something which couldn't be possible. Also, the reason I have chosen a bilingual participant was due to the fact that the original study (Belin et al., 2000) had a sample of English speaking respondents only.
- ix) When we refer to 'non-verbal affect bursts' we mean laughter, screams of fear or not, as well as non-speech sounds of phonemic structure, such as 'wow', 'eh', 'oi', 'ah', etc. (Hester, 2005; Schröder, 2003).
- x) That goes in line with studies having to do with mild and/or severe learning difficulties found the spectrum of autism (Bishop, 2007).

Funding

The author has no funding to report.

Competing Interests

The author has declared that no competing interests exist.

Acknowledgments

The author has no support to report.

References

- Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language*, *24*, 89-106. doi:10.1016/0749-596X(85)90017-8
- Barrett, L., Dunbar, R., & Lycett, J. (2002). *Human evolutionary psychology*. New York, NY: Palgrave.
- Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: Neural correlated of voice perception. *Trends in Cognitive Sciences*, *8*(3), 129-135. doi:10.1016/j.tics.2004.01.008
- Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The montreal affective voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods*, *40*(2), 531-539. doi:10.3758/BRM.40.2.531
- Belin, P., Zatorre, R. J., & Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Cognitive Brain Research*, *13*, 17-26. doi:10.1016/S0926-6410(01)00084-2
- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*, 309-312. doi:10.1038/35002078
- Bélizaire, G., Fillion-Bilodeau, S., Chartrand, J. P., Bertrand-Gauvin, C., & Belin, P. (2007). Cerebral response to 'voiceness': A functional magnetic resonance imaging study. *Neuroreport*, *18*(1), 29-33.
- Benzagmout, M., Chaoui, M. E. F., & Duffau, H. (2008). Reversible deficit affecting the perception of tone of a human voice after tumour resection from the right auditory cortex. *Acta Neurochirurgica*, *150*(6), 589-593. doi:10.1007/s00701-008-1495-4
- Binder, J. R., Frost, J. A., & Bellgowan, P. S. F. (1999). Superior temporal sulcus (STS) responses to speech and nonspeech auditory stimuli. *Journal of Cognitive Neuroscience*, *11*(Supp. 1), 99.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, *10*(5), 512-528. doi:10.1093/cercor/10.5.512
- Binder, J. R., Frost, J. A., Hammeke, T. A., Rao, S. M., & Cox, R. W. (1996). Function of the left planum temporale in auditory and linguistic processing. *Brain*, *119*(4), 1239-1247. doi:10.1093/brain/119.4.1239
- Bishop, D. V. M. (2007). Using mismatch negativity to study central auditory processing in developmental language and literacy impairments: Where are we, and where should we be going? *Psychological Bulletin*, *133*, 651-672. doi:10.1037/0033-2909.133.4.651
- Bishop, D. V. M., & Hardiman, M. J. (2010). Measurement of mismatch negativity in individuals: A study using single-trial analysis. *Psychophysiology*, *47*, 697-705.
- Bonte, M., Valente, G., & Formisano, E. (2009). Dynamic and task-dependent encoding of speech and voice by phase reorganization of cortical oscillations. *The Journal of Neuroscience*, *29*(6), 1699-1706. doi:10.1523/JNEUROSCI.3694-08.2009

- Brown, C. P. (1999). Sex and hemispheric differences for rapid auditory processing in normal adults. *Laterality: Asymmetries of Body, Brain and Cognition*, 4(1), 39-50.
- Buchweitz, A., Mason, R. A., Tomitch, L. M. B., & Just, M. A. (2009). Brain activation for reading and listening comprehension: An fMRI study of modality effects and individual differences in language comprehension. *Psychology & Neuroscience*, 2(2), 111-123. doi:10.3922/j.psns.2009.2.003
- Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, 11(12), 535-543. doi:10.1016/j.tics.2007.10.001
- Campbell, R. (2008). The processing of audio-visual speech: Empirical and neural bases. *Philosophical Transactions of the Royal Society: Series B. Biological Sciences*, 363(1493), 1001-1010. doi:10.1098/rstb.2007.2155
- Charest, I., Pernet, C. R., Rousselet, G. A., Quiñones, I., Latinus, M., Fillion-Bilodeau, S., . . . Belin, P. (2009). Electrophysiological evidence for an early processing of human voices. *BMC Neuroscience*, 10, Article 127.
- Corning, P. A. (2000). Biological adaptation in human societies: A 'basic needs' approach. *Journal of Bioeconomics*, 2, 41-86. doi:10.1023/A:1010027222840
- Cornwell, B. R., Baas, J. M. P., Johnson, L., Holroyd, T., Carver, F. W., Lissek, S., & Grillon, C. (2007). Neural responses to auditory stimulus deviance under threat of electric shock revealed by spatially-filtered magnetoencephalography. *NeuroImage*, 37(1), 282-289. doi:10.1016/j.neuroimage.2007.04.055
- Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., . . . Wildgruber, D. (2006). Impact of voice on emotional judgment of faces: An event-related fMRI study. *Human Brain Mapping*, 27(9), 707-714. doi:10.1002/hbm.20212
- Ethofer, T., Anders, S., Wiethoff, S., Erb, M., Herbert, C., Saur, R., . . . Wildgruber, D. (2006). Effects of prosodic emotional intensity on activation of associative auditory cortex. *Neuroreport*, 17, 249-253. doi:10.1097/01.wnr.0000199466.32036.5d
- Fabbro, F. (2001). The bilingual brain: Cerebral representation of languages. *Brain and Language*, 79, 211-222. doi:10.1006/brln.2001.2481
- Fecteau, S., Armony, J. L., Joanette, Y., & Belin, P. (2004). Is voice processing species-specific in human auditory cortex? An fMRI study. *NeuroImage*, 23, 840-848. doi:10.1016/j.neuroimage.2004.09.019
- Fiske, M. H., & Ofshe, R. (1970). The process of status evolution. *Sociometry*, 33, 327-346. doi:10.2307/2786161
- Ghee, M., Soon, C. S., & Hwee, L. L. (2005). The influence of language experience on cortical activation in bilinguals. In J. Cohen, K. T. McAlister, K. Rolstad, & J. MacSwan (Eds.), *ISB4: Proceedings of the 4th International Symposium on bilingualism* (pp. 522-526). Somerville, MA: Cascadilla Press.
- Grossmann, T., Oberecker, R., Koch, S. P., & Friederici, A. D. (2010). The developmental origins of voice processing in the human brain. *Neuron*, 65, 852-858. doi:10.1016/j.neuron.2010.03.001
- Gygi, B., Kidd, G. R., & Watson, C. S. (2007). Similarity and categorization of environmental sounds. *Perception & Psychophysics*, 69(6), 839-855. doi:10.3758/BF03193921
- Hanley, J. R., Smith, S. T., & Hadfield, J. (1998). "I recognise you but I can't place you": An investigation of familiar-only experiences during tests of voice and face recognition. *The Quarterly Journal of Experimental Psychology: Section A. Human Experimental Psychology*, 51, 179-195. doi:10.1080/713755751

- Hashimoto, R., Homae, F., Nakajima, K., Miyashita, Y., & Sakai, K. L. (2000). Functional differentiation in the human auditory and language areas revealed by a dichotic listening task. *NeuroImage*, *12*, 147-158. doi:10.1006/nimg.2000.0603
- Hester, E. (2005). The evolution of the auditory system: A tutorial. *Contemporary Issues in Communication Science and Disorders*, *32*, 5-10.
- Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*, 528-551. doi:10.1037/0033-295X.95.4.528
- Imaizumi, S., Koichi, M., Shigeru, K., Ryuta, K., Motoaki, S., Hiroshi, F., . . . Katsuki, N. (1997). Vocal identification of speaker and emotion activates different brain regions. *Neuroreport*, *8*(12), 2809-2812. doi:10.1097/00001756-199708180-00031
- Jaramillo, M., Ilvonen, T., Kujala, T., Alku, P., Tervaniemi, M., & Alho, K. (2001). Are different kinds of acoustic features processed differently for speech and non-speech sounds? *Cognitive Brain Research*, *12*, 459-466. doi:10.1016/S0926-6410(01)00081-7
- Joanisse, M. F., & Gati, J. S. (2003). Overlapping neural regions for processing rapid temporal cues in speech and nonspeech signals. *NeuroImage*, *19*, 64-79. doi:10.1016/S1053-8119(03)00046-6
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialised for face perception. *The Journal of Neuroscience*, *17*, 4302-4311.
- Kolb, B., & Whishaw, I. Q. (2008). *Fundamentals of human neuropsychology*. New York, NY: Worth Publishers.
- Krantz, J. (2012). *Experiencing sensation and perception*. New York, NY: Pearson.
- Levy, D. A., Granot, R., & Bentin, S. (2001). Processing specificity for human voice stimuli: Electrophysiological evidence. *NeuroReport*, *12*(12), 2653-2657.
- Locke, J. L. (2008). Cost and complexity: Selection for speech and language. *Journal of Theoretical Biology*, *251*, 640-652. doi:10.1016/j.jtbi.2007.12.022
- Moore, R. E., Estis, J., Gordon-Hickey, S., & Watts, C. (2008). Pitch discrimination and pitch matching abilities with vocal and nonvocal stimuli. *Journal of Voice*, *22*(4), 399-407. doi:10.1016/j.jvoice.2006.10.013
- Morillon, B., Lehongre, K., Frackowiak, R. S. J., Ducorps, A., Kleinschmidt, A., Poeppel, D., & Giraud, A. L. (2010). Neurophysiological origin of human brain asymmetry for speech and language. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(43), 18688-18693. doi:10.1073/pnas.1007189107
- Näätänen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Lawrence Erlbaum.
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The Mismatch Negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*, 2544-2590.
- Palmer, A. R., & Summerfield, A. Q. (2002). Microelectrode and neuroimaging studies of central auditory function. *British Medical Bulletin*, *63*(1), 95-105. doi:10.1093/bmb/63.1.95
- Patterson, R. D., & Johnsrude, I. S. (2008). Functional imaging of the auditory processing applied to speech sounds. *Philosophical Transactions of the Royal Society: Series B. Biological Sciences*, *363*, 1023-1035. doi:10.1098/rstb.2007.2157

- Petkov, C. I., Kayser, C., Steudel, T., Whittingstall, K., Augath, M., & Logothetis, N. K. (2008). A voice region in the monkey brain. *Nature Neuroscience*, *11*, 367-374. doi:10.1038/nn2043
- Price, C. J., Green, D. W., & von Studnitz, R. (1999). A functional imaging study of translation and language switching. *Brain*, *122*(12), 2221-2235. doi:10.1093/brain/122.12.2221
- Rothmayr, C., Baumann, O., Endestad, T., Rutschmann, R. M., Magnussen, S., & Greenlee, M. (2007). Dissociation of neural correlates of verbal and non-verbal visual working memory with different delays. *Behavioral and Brain Functions*, *3*, 56-67. doi:10.1186/1744-9081-3-56
- Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice*, *9*, 235-248. doi:10.1016/S0892-1997(05)80231-0
- Schneider, P., Scheng, M., Dosch, H. G., Specht, H. J., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience*, *5*(7), 688-694. doi:10.1038/nn871
- Schröder, M. (2003). Experimental study of affect bursts. *Speech Communication*, *40*, 99-116. doi:10.1016/S0167-6393(02)00078-X
- Shultz, S., Vouloumanos, A., & Pelphrey, K. (2012). The superior temporal sulcus differentiates communicative and noncommunicative auditory signals. *Journal of Cognitive Neuroscience*, *24*(5), 1224-1232. doi:10.1162/jocn_a_00208
- Steinschneider, M. (2012). Phonemic representations and categories. In Y. E. Cohen, A. N. Popper, & R. R. Fay (Eds.), *Neural correlates of auditory cognition* (pp. 151-192). New York, NY: Springer.
- Sweller, J. (2004). Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. *Instructional Science*, *32*, 9-31. doi:10.1023/B:TRUC.0000021808.72598.4d
- Traxler, M. J. (2012). *Introduction to psycholinguistics: Understanding language science*. Chichester, United Kingdom: John Wiley & Sons.
- Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science*, *311*, 670-674. doi:10.1126/science.1119983
- Uppenkamp, S., Johnsrude, I., Patterson, R. D., Norris, D., & Marslen-Wilson, W. (2006). Locating the initial stages of speech-sound processing in human temporal cortex. *NeuroImage*, *31*, 1284-1296. doi:10.1016/j.neuroimage.2006.01.004
- Varvatsoulias, G. (2013). The physiological processes underpinning PET and fMRI techniques with an emphasis on the temporal and spatial resolution of these methods. *Psychological Thought*, *6*(2), 173-195. doi:10.5964/psyc.v6i2.75
- von Kriegstein, K., Eger, E., Kleinschmidt, A., & Giraud, A. L. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Cognitive Brain Research*, *17*, 48-55. doi:10.1016/S0926-6410(03)00079-X
- Vouloumanos, A., Kiehl, K. A., Werker, J. F., & Liddle, P. F. (2001). Detection of sounds in the auditory stream: Event-related fMRI evidence for differential activation to speech and nonspeech. *Journal of Cognitive Neuroscience*, *13*(7), 994-1005. doi:10.1162/089892901753165890
- Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*, 585-594. doi:10.1016/j.tics.2005.10.011

- Ward, J. (2010). *The student's guide to cognitive neuroscience*. Hove, United Kingdom: Psychology Press.
- Westman, J. C., & Walters, J. R. (1981). Noise and stress: A comprehensive approach. *Environmental Health Perspectives*, *41*, 291-309. doi:10.1289/ehp.8141291
- Whalen, D. H., Benson, R. R., Richardson, M., Swainson, B., Clark, V. P., Lai, S., . . . Liberman, A. M. (2006). Differentiation of speech and nonspeech processing within primary auditory cortex. *The Journal of the Acoustical Society of America*, *119*(1), 575-581. doi:10.1121/1.2139627
- Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(20), 11125-11129. doi:10.1073/pnas.97.20.11125
- Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, *256*, 846-849. doi:10.1126/science.1589767

About the Author

Dr. **George Varvatsoulis** CPsychol CSci Expert Witness, studied Theology and Practical Theology in Thessaloniki/Greece. In England, he studied Psychology, Psychology of Religion, Evolutionary Psychology and Cognitive-Behavioural Therapy as well as he holds a Doctorate in Psychology of Religion from the University of Durham (Northeastern England). He has published 49 articles in peer-reviewed Journals and three books. He had worked as a Cognitive-Behavioural Practitioner at the West London Mental Health NHS (National Health Service) Trust where he was seeing clients with co-morbid conditions. He has also been a Visiting Research Fellow for the topic of Psychology of Religion at Glyndŵr University in Wrexham/Wales, UK. At present, he is a Lecturer at Newham University Centre, East London, teaching cognitive-behavioural therapy (CBT), research methods and integrative counselling in the BA/BSc Combined/Counselling Studies (CBT & Integrative) Programme.