What do we know about hearing and brain function related to voice? A review

Mette Pedersen, Ear Nose specialist, FRSM PhD et h.c. research prof. IBC Cambridge, UK. M. Shezad Mahmood, Sanila Mahmood, Bilal Hussain Akram, Anders Jønsson, Anne A. Agersted. University of Copenhagen.

Abstract

This review examines what we know about hearing and brain function related to voice. The aim was to investigate the role of the auditory cortex in the processing of speech, songs and other vocalizations, which all are spectrally and temporally highly structured sounds. A British Library search was made. For the latest 5 years 812 references were found. Few relevant articles were found on the hearing process related to voice. This review presents the research made on the subject till now. This search shows that the area has some interest but further research is important.

Introduction

Basic anatomical structure overview of the ear

Anatomically the ear is divided in three parts, the outer ear, middle ear and inner ear. Sound waves are collected by the earlobe which is a part of the outer ear, the collected waves travel along the ear canal and make the membrana tympani move as located between the outer and the middle ear. The membrana tympani vibrations stimulate the small bones of the ear, which amplify the waves. The outer haircells provide a feed-back reinforcement of vibration of the basilar membrane and make cochlear sensitive. Cochlea turns the waves into electrical signals, and these signals are transported by the auditory nerve to the brain. Nerve signals propagate with a rate of approximately 10-30m/s and will therefore be delayed up through the hearing way. In each synapse is a further delay of approximately 1ms, figure 1.

Fig. 1 (ref.1)

Model of neural processing in cortex

Communication calls consist of elementary features, such as bandpass noise bursts or frequency-modulated (FM) sweeps, figure 2 (2). Harmonic calls, such as the vocal scream from the rhesus monkey repertoire depicted here by its spectrogram and time signal amplitude (A, measured as output voltage of a sound meter), consist of fundamental frequencies and higher harmonics.
The neural circuitry for processing such calls is thought to consist of small hierarchical networks. At the lowest level, there are neurons serving as FM detectors tuned to the rate and direction of FM sweeps; these detectors extract each FM component (shown in cartoon spectrograms) in the upward and downward sweeps of the scream. The output of these FM detectors is combined nonlinearly at the next level: the target neurons T1 and T2 possess a high threshold and fire only if all inputs are activated. At the final level, a 'tonal-scream detector' is created by again combining output from neurons T1 and T2 nonlinearly. Temporal integration is accomplished by having the output of T1 pass through a delay line with a latency $\Delta t1$ sufficient to hold up the input to the top neuron long enough that all inputs arrive at the same time (2).

Early processing of human speech sounds in the antero-lateral auditory belt and parabelt cortex is thought to be accomplished in a similar way.
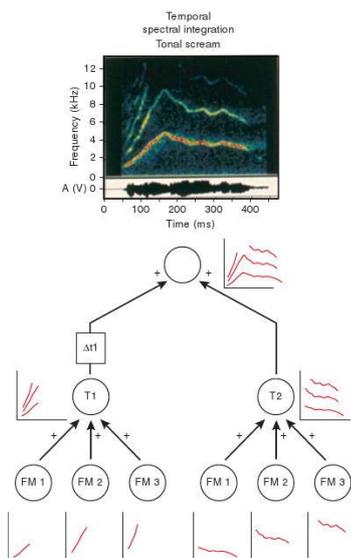

Fig. 2 (2).

Main brain regions involved in auditory processing

Dual auditory processing scheme of the human brain and the role of internal models in sensory systems is shown in figure 3 (3). This expanded scheme closes the loop between speech perception and production and proposes a common computational structure for space processing and speech control in the postero-dorsal auditory stream.

(a) Antero-ventral (green) and postero-dorsal (red) streams originating from the auditory belt. The postero-dorsal stream interfaces with premotor areas and pivots around the inferior parietal lobe (IPL) where a quick sketch of sensory event information is compared with a predictive efference copy of motor plans.

(c) In one direction, the model performs a forward mapping: object information, such as speech, is decoded in the antero-ventral stream all the way to category-invariant inferior frontal cortex (IFC) (area 45), and is transformed into motor-articulatory representations (area 44 and ventral PMC), whose activation is transmitted to the IPL (and posterior superior temporal cortex) as an efference copy. (b) In reverse direction, the model performs an inverse mapping, whereby attention- or intention-related changes in the IPL influence the selection of context-dependent action programs in prefrontal cortex (PFC) and premotor cortex (PMC).

Both types of dynamic model are testable using techniques with high temporal precision (for example, magnetoencephalography in humans or single-unit studies in monkeys) that allow determination of the order of events in the respective neural systems (3)
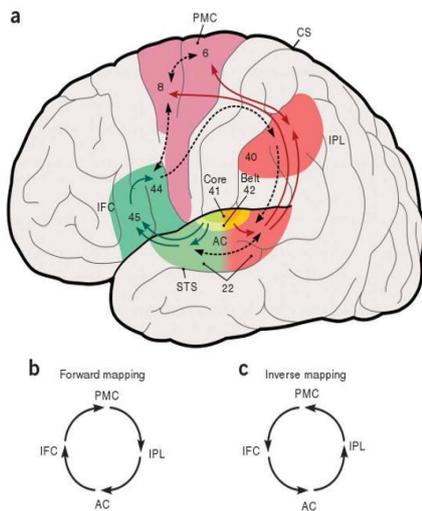

Fig. 3(4)

Overview of auditory language processing of PET studies

Figure 4 depicts the impact of several factors on the distribution of voice and language-related neural activities. Schematic activations are redrawn from the original studies using the significance threshold reported in each study.

- *A*: Rapid acoustic transitions elicit more activity in the left superior temporal gyri (STG) than contralateral right-sided regions.

- *B* and *C*: Increasing difficulty of phoneme monitoring tasks induces a left-sided asymmetry in the activity of the STG.

- *D*: Increased activity in the dorsal pathway by contrasting phoneme monitoring (the difficult variant) with semantic categorization (the reverse contrast is shown in *J*).

- *E*: Intelligible speech samples activate a specific pathway along the anterior part of the left superior temporal sulcus. Pale red depicts activations elicited by speech like stimuli including unintelligible samples, and red depicts intelligible speech activations. The white cluster in the right hemisphere is interpreted as a correlate of dynamic pitch perception.

- *F* and *G*: Two regions specifically activated for accessing semantic contents from spoken words vs. environmental sounds (anterior part of the left STG) (*F*) and environmental sounds vs. spoken words (posterior part of the right STG) (*G*). Note the congruence of *E, left*, and *F*. *H*: Common activations for verbal and nonverbal meaningful inputs relative to matched noises in perisylvian regions.
I and *J*: A set of regions activated during semantic processing compared with pure tone monitoring (*I*) or phoneme detection (*J*). In addition to the activity in the anterior part of the left superior temporal sulcus, the "ventral" pathway involved in lexical semantic processing included an area at the junction left inferior and middle temporal gyri. The right angular gyrus appears activated in *J* and not in *I* because it was activated to the same level in semantic categorization and pure tone monitoring in *I*. (3).
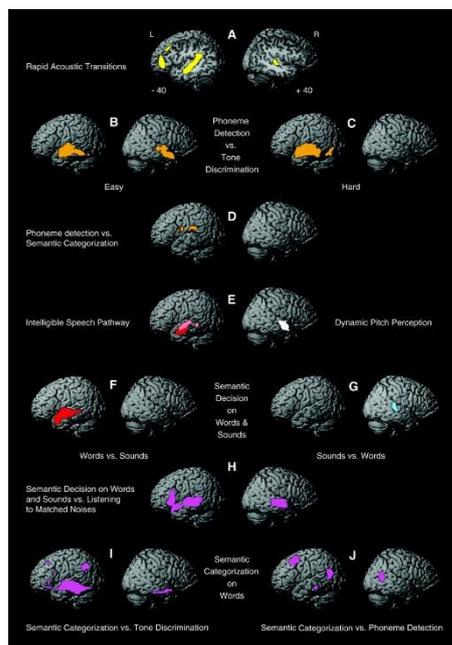
Why no study on singing?

Fig. 4(3)

Understanding voice perception

Voices carry large amounts of socially relevant information on persons, much like 'auditory faces'. Following *Bruce and Young (1986)*'s seminal model of face perception (5), it is proposed that the

cerebral processing of vocal information is organized in interacting but functionally dissociable pathways for processing the three main types of vocal information: speech, identity, and affect. The predictions of the 'auditory face' model of voice perception is reviewed in the light of recent clinical, psychological, and neuroimaging evidence (6).

Human voice perception

We are all voice experts. First and foremost, we can produce and understand speech, and this makes us a unique species. But in addition to speech perception, we routinely extract from voices a wealth of socially-relevant information in what constitutes a more primitive, and probably more universal, non-linguistic mode of communication.

Consider the following example: you are sitting in a plane, and you can hear a conversation in a foreign language in the row behind you.

You do not see the speakers' faces, and you cannot understand the speech content because you do not know the language. Yet, an amazing amount of information is available to you. You can evaluate the physical characteristics of the different protagonists, including their gender, approximate age and size, and associate an identity to the different voices. You can form a good idea of the different speaker's mood and affective state, as well as more subtle cues as the perceived attractiveness or dominance of the protagonists. In brief, you can form a fairly detailed picture of the type of social interaction unfolding, which a brief glance backwards can on the occasion help refine — sometimes surprisingly so.

What are the acoustical cues that carry these different types of vocal information? How does our brain process and analyze this information? Briefly review is made of an emerging field and the main tools used in voice perception research (7).

How do we recognize who is speaking?

The human brain effortlessly extracts a wealth of information from natural speech, which allows the listener to both understand the speech message and recognize who is speaking.

This article reviews behavioral and neuroscientific work that has attempted to characterise how listeners achieve speaker recognition. Behavioral studies suggest that the action of a speaker's glottal folds and the overall length of their vocal tract carry important voice-quality information. Although these cues are useful for discriminating and recognising speakers under certain circumstances, listeners may use virtually any systematic feature for recognition. Neuroscientific studies have revealed that speaker recognition relies upon a predominantly right-lateralised network of brain regions.

Specifically, the posterior parts of superior temporal sulcus appear to perform some of the acoustical analyses necessary for the perception of speaker and message, whilst anterior portions may play a more abstract role in perceiving speaker identity.

This voice-processing network is supported by direct, early connections to non-auditory regions, such as the visual face-sensitive area in the fusiform gyrus, which may serve to optimize person recognition. (8)

How does the auditory cortex neurons represent communication sounds?

The present review aims at investigating the role of auditory cortex in the processing of speech, bird songs and other vocalizations, which all are spectrally and temporally highly structured sounds.

Whereas earlier studies have simply looked for neurons exhibiting higher firing rates to conspecific vocalizations over their modified, artificially synthesized versions, more recent studies determined the coding capacity of temporal spike patterns, which are prominent in primary and non-primary areas (and also in non-auditory cortical areas).

In several cases, this information seems to be correlated with the behavioral performance of human or animal subjects, suggesting that spike-timing based coding strategies might set the foundations of our perceptive abilities. Also, it is now clear that the responses of auditory cortex neurons are highly nonlinear and that their responses to natural stimuli cannot be predicted from their responses to artificial stimuli such as moving ripples and broadband noises.

The auditory cortex neurons cannot follow rapid fluctuations of the vocalizations envelope, they only respond at specific time points during communication sounds, which can serve as temporal markers for integrating the temporal and spectral processing taking place at subcortical relays.

Thus, the temporal sparse code of auditory cortex neurons can be considered as a first step for generating high level representations of communication sounds independent of the acoustic characteristic of these sounds. (9).

Representation of speech in human auditory cortex: is it special?

Successful categorization of phonemes in speech requires that the brain analyze the acoustic signal along both spectral and temporal dimensions. Neural encoding of the stimulus amplitude envelope is critical for parsing the speech stream into syllabic units.

Encoding of voice onset time (VOT) and place of articulation (POA), cues necessary for determining phonemic identity, occurs within shorter time frames. An unresolved question is whether the neural representation of speech is based on processing mechanisms that are unique to humans and shaped by learning and experience, or is based on rules governing general auditory processing that are also present in non-human animals. This question was examined by comparing the neural activity elicited by speech and other complex vocalizations in primary auditory cortex of macaques, who are limited vocal learners, with that in Heschl's gyrus, the putative location of primary auditory cortex in humans. Entrainment to the amplitude envelope is neither specific to humans nor to human speech.

VOT is represented by responses time-locked to consonant release and voicing onset in both humans and monkeys. Temporal representation of VOT is observed both for isolated syllables and for syllables embedded in the more naturalistic context of running speech.

The fundamental frequency of male speakers is represented by more rapid neural activity phase-locked to the glottal pulsation rate in both humans and monkeys.
In both species, the differential representation of stop consonants varying in their POA can be predicted by the relationship between the frequency selectivity of neurons and the onset spectra of the speech sounds.

These findings indicate that the neurophysiology of primary auditory cortex is similar in monkeys and humans despite their vastly different experience with human speech, and that Heschl's gyrus is engaged in general auditory, and not language-specific, processing (10).

Auditory signal processing in communication: perception and performance of vocal sounds

Learning and maintaining the sounds we use in vocal communication require accurate perception of the sounds we hear performed by others and feedback-dependent imitation of those sounds to produce our own vocalizations. Understanding how the central nervous system integrates auditory and vocal-motor information to enable communication is a fundamental goal of systems neuroscience, and insights into the mechanisms of those processes will profoundly enhance clinical therapies for communication disorders.

Gaining the high-resolution insight necessary to define the circuits and cellular mechanisms underlying human vocal communication is presently impractical. Songbirds are the best animal model of human speech, and this review highlights recent insights into the neural basis of auditory perception and feedback-dependent imitation in those animals. Neural correlates of song perception are present in auditory areas, and those correlates are preserved in the auditory responses of downstream neurons that are also active when the bird sings. Initial tests indicate that singing-related activity in those downstream neurons is associated with vocal-motor performance as opposed to the bird simply hearing itself sing. Therefore, action potentials related to auditory perception and action potentials related to vocal performance are co-localized in individual neurons.

Conceptual models of song learning involve comparison of vocal commands and the associated auditory feedback to compute an error signal that is used to guide refinement of subsequent song performances, yet the sites of that comparison remain unknown. Convergence of sensory and motor activity onto individual neurons points to a possible mechanism through which auditory and vocal-motor signals may be linked to enable learning and maintenance of the sounds used in vocal communication (11).

Methods

A literature search at British Library London has been made. 812 references were found with the search words: Hearing AND voice AND brain function – for the latest 5 years. This shows that the area has some scientific interest.

Aspects of pathology 1

Hearing loss in singers: a preliminary study

Singers need good hearing; however, they may be exposed to loud noises during their musical activities. The objectives of this study were to describe the incidence and type of hearing loss (HL) in singers. A retrospective case cohort study was made (12).

Billing records identified patients who had undergone videostroboscopy and audiogram during the same visit over a 3 year period. A **singer** was defined as anyone who self-identified as a singer (professional or avocational). Age and gender matched non singers were used as controls. Patients with otologic diagnoses, surgery, or complaints were excluded. Retrospective chart review was conducted for the presence of HL, type of HL, and pure tones audiogram results. Statistical analysis included descriptive statistics, Students t test, chi-square test, and Fisher exact test.

Of 172 singers (44.7 years, 37.8% male), 31 (17.5%) had hearing loss (HL). Pure tone thresholds for the singers with HL subgroup at 3, 4, and 6 kHz were 21.0, 26.5, and 34.4 dB in the right, and 22.8, 30.3, and 38.8 dB in the left ear, respectively. Older age, male gender, longer number of years of singing and baritone voice were associated with HL. There was no association with genre of music. When compared with controls, the incidence of HL (19.8%) was not significantly different. Pure tones at 3, 4, and 6 kHz were not significantly different than controls with HL. Most common type of HL in singers was bilateral sensorineural (83.9%), which was significantly higher than controls .

The incidence of HL in singers was 17.5%, which was not significantly different from controls. Bilateral sensorineural HL was most common (12).

Variability in voice fundamental frequency of sustained vowels in speakers with sensorineural hearing loss

In a previous study, the low-frequency modulation extent (LFP) of the vocal fundamental frequency (F(0)) showed a significant increase in the presence of binaural noise masking for the healthy individuals. This study was to investigate the F(0) of subjects with sensorineural hearing loss (SNHL) using sustained phonations to explore the changes of F(0) modulations in SNHL. Twenty-three SNHL subjects and 14 age-matched subjects without hearing loss were enrolled in the study. Sustained vocalizations of vowel /a/ for more than 5 seconds were digitally recorded. The F(0) contour of each phonation was acquired using digital signal processing. The modulation extent at different frequencies was obtained using Fourier transformation of F(0) contour. The LFP of F(0) (<3Hz) was significantly greater for the SNHL subjects (P<0.001, independent samples t test). Although the correlation analysis was limited to the auditory-evoked brainstem response

(ABR) thresholds because of their disagreement with the pure-tone thresholds in some subjects with functional hearing disorder, the correlation between LFP and ABR thresholds was significant (ρ=0.45, P=0.03, Spearman's correlation analysis).

The LFPs of F(0) were significantly greater for the SNHL subjects and the changes of F(0) modulations could be detected using power spectral analysis of F(0). The method may be used for evaluation of audio-vocal feedback in SNHL (13).

Aspects of pathology 2

Normal-hearing listeners' and cochlear implant users' perception of pitch cues in emotional speech

In cochlear implants (CIs), acoustic speech cues, especially for pitch, are delivered in a degraded form. This study's aim was to assess whether due to degraded pitch cues, normal-hearing listeners and CI users employ different perceptual strategies to recognize vocal emotions, and, if so, how these differ. Voice actors were recorded pronouncing a nonce word in four different emotions: anger, sadness, joy, and relief. These recordings' pitch cues were phonetically analyzed. The recordings were used to test 20 normal-hearing listeners' and 20 CI users' emotion recognition.

Normal-hearing listeners outperformed CI users in emotion recognition, even when presented with CI simulated stimuli. However, only normal-hearing listeners recognized one particular actor's emotions worse than the other actors'. The groups behaved differently when presented with similar input, showing that they had to employ differing strategies.
Considering the respective speaker's deviating pronunciation, it appears that for normal-hearing listeners, mean pitch is a more salient cue than pitch range, whereas CI users are biased toward pitch range cues (14).

Systemic analysis of the benefits of cochlear implants on voice production

The voice of hearing impaired children and adults with cochlear implants has been little studied. There is not an effective number of studies with high evidence levels which precisely show the effects of the cochlear implantation on the quality of voice of these individuals (15).

Perception of music timbre by cochlear implant listeners: a multidimensional scaling study

Several studies have shown that the ability to identify the timbre of musical instruments is reduced in cochlear implant (CI) users compared with normal-hearing (NH) listeners. However, most of these studies have focused on tasks that require specific musical knowledge.

In contrast, the present study investigates the perception of timbre by CI subjects using a multidimensional scaling (MDS) paradigm. The main objective was to investigate whether CI subjects use the same cues as NH listeners do to differentiate the timbre of musical instruments. The internal representation of musical timbre for isolated musical instrument sounds was found to be similar in NH and in CI listeners. This suggests that training procedures designed to improve

timbre recognition in CIs will indeed train CI subjects to use the same cues as NH listeners. Furthermore, NH subjects listening to noise-vocoded sounds appear to be a good model of CI timbre perception as they show the same first two perceptual dimensions as CI subjects do and also exhibit a similar change in perceptual weights applied to these two dimensions. This last finding validates the use of simulations to evaluate and compare training procedures to improve timbre perception in Cis (16).

Auditory cortex activation to natural speech and stimulated cochlear implant speech measured with functional near-infrared spectroscopy

This study determined whether functional near-infrared spectroscopy (fNIRS), a non-invasive neuroimaging method which is safe to use repeatedly and for extended periods of time, can provide an objective measure of whether a subject is hearing normal speech or distorted speech. A 140 channel fNIRS system were used to measure activation within the auditory cortex in 19 normal hearing subjects while they listed to speech with different levels of intelligibility. Custom software was developed to analyze the data and compute topographic maps from the measured changes in oxyhemoglobin and deoxyhemoglobin concentration.
Normal speech reliably evoked the strongest responses within the auditory cortex. Distorted speech produced less region-specific cortical activation.

Environmental sounds were used as a control, and they produced the least cortical activation. These data collected using fNIRS are consistent with the fMRI literature and thus demonstrate the feasibility of using this technique to objectively detect differences in cortical responses to speech of different intelligibility (17).

The use of acoustic cues for phonetic identification: effects of spectral degradation and electric hearing

Although some cochlear implant (CI) listeners can show good word recognition accuracy, it is not clear how they perceive and use the various acoustic cues that contribute to phonetic perceptions. In this study, the use of acoustic cues was assessed for normal-hearing (NH) listeners in optimal and spectrally degraded conditions, and also for CI listeners. Two experiments tested the tense/lax vowel contrast (varying in formant structure, vowel-inherent spectral change, and vowel duration) and the word-final fricative voicing contrast (varying in F1 transition, vowel duration, consonant duration, and consonant voicing).

The results suggest that although NH and CI listeners may show similar accuracy on basic tests of word, phoneme or feature recognition, they may be using different perceptual strategies in the process (18).

Aspects of pathology 3

The role of fundamental frequency and formants in voice gender identification
F0 was found to be the primary cue for gender perception and listeners showed a higher accuracy

in identifying male's than female's voices. The findings are consistent with results previously reported, although other acoustic cues such as voice quality may also affect gender perception (19).

Does knowledge of medical diagnosis bias auditory-perceptual judgments of dysphonia
Regardless of experience level, listeners judged the samples as significantly more severe when the speakers' diagnoses were known. Specifically, novice listeners (NLs) significantly increased the severity of judgments for speakers who were mildly breathy or mildly or moderately rough when diagnostic information was known (20).

Functional magnetic resonance imaging study of brain activity associated with pitch adaptation during phonation in healthy women without voice disorders
This study demonstrated that a blocked design fMRI is sensitive enough to define a widespread network of activation associated with phonation involving pitch variation. The results of this study will be implemented in our future research on phonation and its disorders (21).

Voice outcomes of laryngopharyngeal reflux treatment: a systemic review of 1483 patients

The main voice assessment outcomes reported were hoarseness assessments by physicians or patients, followed by acoustic parameters; 15 and 14 articles, respectively, demonstrated significant improvements in subjective and objective voice assessments after treatment.
The methodology used to measure acoustic parameters (i.e. sustained vowel duration, the sample portion choice for measurement, etc.) varied from one study to another. The majority of studies indicated that voice quality assessments (especially acoustic parameters) remain an interesting outcome to measure the effectiveness of treatment, but further studies using standardised and transparent methodology to measure acoustic parameters are necessary to confirm the place of each tool in the LPR disease evaluation (22).

Relationship between acoustic measurements and self-evaluation in patients with voice disorders
There was a correlation between the scores of voice symptom scale (VoiSS) and acoustic measurements. Patients with self-reported voice problems in VoiSS present greater deviations in acoustic measures, mainly in jitter. There was no correlation between the Voice handicap index (VHI) scores and the acoustic measures and no difference in the averages of these measures between patients with and without voice problems detected from the VHI cutoffs (23).

Vocal performance of group fitness instructors before and after instruction: changes in acoustic measures and self-ratings

To quantify acute changes in acoustic parameters of the voices of group fitness instructors (GFIs) before and after exercise instruction. Significant increases ($P \leq 0.05$) were found in fundamental frequency and intensity after instruction. No significant changes in jitter, shimmer, HNR, or MDSP were found before and after instruction. For the group, no significant change in self-ratings of voice quality occurred before and after instruction (24).

Subcortical modulation in auditory processing and auditory hallucinations

Patients' auditory hallucination severity was assessed by the auditory hallucination item of the Brief Psychiatric Rating Scale (BPRS). During perception of acoustically reversed English (arEW), bilateral activation of the globus pallidus correlated with severity of auditory hallucinations. EW when compared with words (arEW) did not correlate with auditory hallucination severity. Our findings suggest that the sensitivity of the globus pallidus to the human voice is associated with the severity of auditory hallucination (25).

Objective measurement of high level auditory cortical function in children

N2 latency of the cortical auditory evoked potential (CAEP) can be used as an objective measure of temporal processing ability in a paediatric population with temporal processing disorder who are difficult to assess via behavioural response (26).

Voice emotion recognition by cochlear-implanted children and their normally-hearing peers

These results point to the considerable benefit obtained by cochlear-implanted children from their devices, but also underscore the need for further research and development in this important and neglected area (27).

Acoustic properties of vocal singing in prelingually-deafened children with cochlear implants or hearing aids

There is a marked deficit in vocal singing ability either in pitch or rhythm accuracy in a majority of prelingually-deafened children who have received CIs or fitted with HAs. Although an increased length of device use might facilitate singing performance to some extent, the chance for the hearing-impaired children fitted with either HAs or CIs to reach high proficiency in singing is quite slim (28).

Auditory perceptual analysis of voice in abused children and adolescents

This study aimed to analyze the prevalence of vocal changes in abused children and adolescents. Through auditory-perceptual analysis of voice and the study of the association between vocal changes, communication disorders, psychiatric disorders, and global functioning. The prevalence of vocal change was greater than that observed in general population, with significant associations with communication disorders and global functioning. The results demonstrate that the situations these children experience can intensify the triggering of abusive vocal behaviors and consequently, of vocal changes (29).

Conclusion and perspective

This review examines what we know about hearing and brain function related to voice. The aim was to investigate the role of the auditory cortex in the processing of speech, songs and other vocalizations, which all are spectrally and temporally highly structured sounds. A British Library search was made. For the latest 5 years 812 references were found. Few relevant articles were found on the hearing process related to voice. This review presents the research made on the

subject till now. This search shows that the area has some interest but further research is important. The studies of voices were limited in cochlear implants patients but also in many other cases. It is possible to make updated analysis of voice including quantitative measures on vocal fold movements as well as airflow. This means that we have sufficient possibilities to compare hearing, brain function and voice in the future (30-32).

References

(1) http://www.audiologyspecialists.com/anatomy-of-the-ear 09-02-2017
(2) Rauschecker J, Scott S. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. 2009. Nat Neurosci. Vol. 12(6), pp. 718-724
(3) Démont J, Thierry G, Cardebat D. Renewal of the neurophysiology of language: Functional Neuroimaging physiol Rev 2005. Vol. 85, pp. 49-95.
(4) AC, auditory cortex; STS, superior temporal sulcus; IFC, inferior frontal cortex, PMC, premotor cortex; IPL, inferior parietal lobule; CS, central sulcus. Numbers correspond to Brodmann areas.
(5) Bruce V, Young A. Understanding face recognition. 1986. British J. of Psycho. Vol77, pp.305-327.
(6) Belin P, Bestelmeyer P. E. G, Latinus M, Watson R. Understanding voice perception. 2011. British Journal of Psychology
(7) Latinus M, Belin P. Human voice perception. 2011. Curr Biol. Vol.21(4), pp.143-5.
(8) Mathias SR, Von Kriegstein K. How do the recognize who is speaking? 2014. Front Biosci. Vol. 1(6), pp. 92-109.
(9) Gaucher Q, Huetz C, Gourévitch B, Laudanski J, Occelli F, Edeline JM. How do auditory cortex neurons represent communication sounds? 2013. Hear res. Vol. 305, pp. 102-12.
(10) Steinschneider M, Nourski KV, Fishman YI. Representation of speech in human auditory cortex: is it special? 2013. Hear Res. Vol. 305, pp.57-73.
(11) Prather Jf. Auditory signal processing in communication: perception and performances of vocal sounds. 2013. Hear Res. Vol. 305, pp.144-55.
(12) Hu A, Hofmann E, Davis J, Capo J, Krane N, Sataloff RT. Hearing loss in singers: a preliminary study. 2015. J Voice. Vol.29(1), pp.120-4.)
(13) Lee GS. Variability in voice fundamental frequency of sustained vowels in speakers with sensorineural hearing loss. 2012. J Voice. Vol.26(1), pp. 24-9.
(14) Gilbers S, Fuller C, Gilbers D, Broersma M, Goudbeek M, Free R, Baskent D. Normal-hearing listeners' and cochlear implant users' perception of pitch cues in emotional speech. 2016. i-Perception.Vol. 6(5).
(15) Coelho AC, Brasolotto AG, Bevilacqua MC. Systemic analysis of the benefits of cochlear implants on voice production. 2012. J Soc Bras Fonoaudiol. Vol. 24(4), pp.395-402.
(16) Macherey O, Delpierre A. Perception of musical timbre by cochlear implant listerners: a multidimensional scaling study. 2013. Ear Hear. Vol.34(4), pp. 426-36.

(17)   Pollonini L, Olds C, Abaya H, Bortfeld H, Beauchamp M. S, Oghalai J. S. Auditory cortex activation to natural speech and stimulated cochlear implant speech measured with functional near-infrared spectroscopy. 2013. Hearing Research.

(18)   Winn MB, Chatterjee M, Idsardi WJ. The use of acoustic cues for phonetic identification: effects of spectral degradation and electric hearing. 2012. J Acoust Soc Am. Vol.131(2), pp. 1465-79.

(19)   Morris S. F. P, Manwa L. N. The role of fundamental frequency and formants in voice gender identification. 2014. Taylor&Francis. Pp.161-165

(20)   Eadie T, Sroka A, Wright DR, Merati A. Does knowledge of medical diagnosis bias auditory-perceptual judgments of dysphonia? 2011. J Voice. Vol.25(4), pp. 420-9

(21)   Kryshtopava M, Lierde K. V, Meerschman I, D'Haeseleer E, De Moor M, Vandemaele P, Vingerhoets G, Claeys S. Functional magnetic resonance imaging study < of brain activity asiciated with piycj adaptation during phonation healthy women withous voice disorders. 2016. J Voice. Vol.28(6), pp.688-93.

(22)   Lechien JR, Finck C, Costa de Araujo P, Huet K, Delvaux V, Piccaluga Nm Harmegnies B, Saussez S. Voice outcomes of laryngopharyngeal reflux treatment: a systemic review of 1483 patients. 2016. Eur Arch Otorhinoaryngol. Vol.273, pp. 1-23.

(23)   Lopes, L. W, Silva J. D, Simoes L. B, Evangelista D. S, Silva P. O. C, Almeida A. A, Lima-Silva M. F. Relationship between acoustic measurements and self-evaluation in patients with voice disorders. 2016. J Voice.

(24)   Dallaston K, Rumbach AF. Vocal performance of group fitness instructors before and after instruction: changes in acoustic measure and self-ratings. 2016. J Voice. Vol.30(1),pp.127-8.

(25)   Ikuta T, DeRosse P, Argyelan M, Karlsgodt KH, Kingsley PB, Szeszko PR, Malhotra AK. Subcortical modulation in auditory processing and auditory hallucinations. 2015. Behav Brain Res. Vol. 295, pp.78-81.

(26)   Almeqbel A, McMahon C. Objective measurement of high-level auditory cortical function in children. 2015. Int J pediatr Otorhinolaryngol. Vol. 79(7), pp. 1055-62.

(27)   Chatterjee M, Zion DJ, Deroche ML, Burianek BA, Limb CJ, Goran AP, Kulkarni AM, Christensen JA. Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. 2015. Hear res. Vol.322, pp. 151-62.

(28)   Mao Y, Zhang M, Nutter H, Zhang Y, Zhou Q, Liu Q, Wu W, Xie D, Xu L. Acoustica properties of vocal singing in prelingually-deafened children with cochlear implants or hearing aids. 2013. Int J pediatr otorhinolaryngol. Vol.77(11), pp. 1833-40.

(29)   Stivanin L, Santos FP, Oliveira CC, Santos Bd, Ribeiro ST, Scivoletto S. Auditory perceptual analysis of voice in abused children and adolescents. 2015. Braz J Otorhinolaryngology. Vol. 81(1), pp. 71-8.

(30)   Pedersen M (2015) Aspects of resonance: Comparison of high speed films and overtone measurements, Journal of Health Education Research & Development, 3:3 pp. 14

(31)   Pedersen M, Eeg M, Jønsson A, Akram BH. 2015. High-speed films of pathological vocal folds and stiffness of the vocal folds, Proceedings of the 10th Pacific Symposium on Flow Visualization and Image Processing. pp.1-4.

(32)   Pedersen M. 2016. Functional Examination of Voice, a Review, Health Science Journal, Vol. 10(4).