

An Investigation of Male and Female Voices: Does Voice Gender Categorization Depend on Pitch?

Leila Neisi^{1*}, Farahnaz Ahmadpour karimAbadi² & Mansoore Shekaramiz³

^{1,2}Department of Linguistic, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran

³Assistant Professor, Department of Linguistics, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran

Corresponding Author: Leila Neisi, E-mail: leila_neisi@yahoo.com

ARTICLE INFO

Received: September 24, 2019

Accepted: October 27, 2019

Published: November 30, 2019

Volume: 2

Issue: 6

DOI: 10.32996/ijllt.2019.2.6.8

KEYWORDS

*Gender, Voice, Pitch,
Frequency*

ABSTRACT

Gender is a remarkable, socially basic concept got from appearances and voices, yet the cerebrum forms hidden gender segregation have not been greatly investigated. The current study investigated male and female voices and the difference between their pitch. To fulfill, 26 male (n=23) and female (n=23) advance EFL learners were selected from a private English language institute. Firstly, the measured brain response differences were observed between female and male voices beginning at 93 ms. Then, using normal, high- and low-pitch voices, pitch differences was checked between both genders. The findings revealed early effects (31–74 ms) was made by differences in pitch. Gender impacts were viewed with implicit pitch processing, but were not viewed with utilization of pitch. Moreover, stage between 191 and 276 ms differentiated male from female voices, independent of pitch. Thus, these data indicated that voice gender processing included two phases; a very early pitch or frequency distinction and a later more precise distinction of gender.

1. INTRODUCTION

The human voice consists of sound made by a human being using the vocal tract, such as talking, singing, laughing, crying, screaming, shouting, yelling etc. The human voice frequency is specifically a part of human sound production in which the vocal folds (vocal cords) are the primary sound source. (Other sound production mechanisms produced from the same general area of the body involve the production of unvoiced consonants, clicks, whistling and whispering.)

Generally speaking, the mechanism for generating the human voice can be subdivided into three parts; the lungs, the vocal folds within the larynx (voice box), and the articulators. The lungs, the "pump" must produce adequate airflow and air pressure to vibrate vocal folds. The vocal folds (vocal cords) then vibrate to use airflow from the lungs to create audible pulses that form the laryngeal sound source. The muscles of the larynx adjust the length and tension of the vocal folds to 'fine-tune' pitch and tone. The articulators (the parts of the vocal tract above the larynx consisting of tongue, palate, cheek, lips, etc.) articulate and filter the sound emanating from the larynx and to some degree

can interact with the laryngeal airflow to strengthen or weaken it as a sound source.

Adult men and women typically have different sizes of vocal fold; reflecting the male-female differences in larynx size. Adult male voices are usually lower-pitched and have larger folds. The male vocal folds (which would be measured vertically in the opposite diagram), are between 17 mm and 25 mm in length. The female vocal folds are between 12.5 mm and 17.5 mm in length.

The difference in vocal folds size between men and women means that they have differently pitched voices. Additionally, genetics also causes variances amongst the same gender, with men's and women's singing voices being categorized into types. For example, among men, there are bass, baritone, tenor and countertenor (ranging from E2 to even C6 and higher), and among women, contralto, mezzo-soprano and soprano (ranging from F3 to C6 and higher). There are additional categories for operatic voices, see voice type. This is not the only source of difference between male and female voice. Men, generally speaking, have a larger vocal tract, which essentially gives the

resultant voice a lower-sounding timbre. This is mostly independent of the vocal folds themselves.

Human spoken language makes use of the ability of almost all people in a given society to dynamically modulate certain parameters of the laryngeal voice source in a consistent manner. The most important communicative, or phonetic, parameters are the voice pitch (determined by the vibratory frequency of the vocal folds) and the degree of separation of the vocal folds, referred to as vocal fold adduction (coming together) or abduction (separating). The ability to vary the ab/adduction of the vocal folds quickly has a strong genetic component, since vocal fold adduction has a life-preserving function in keeping food from passing into the lungs, in addition to the covering action of the epiglottis. Consequently, the muscles that control this action are among the fastest in the body. Children can learn to use this action consistently during speech at an early age, as they learn to speak the difference between utterances. Surprisingly enough, they can learn to do this well before the age of two by listening only to the voices of adults around them who have voices much different from their own, and even though the laryngeal movements causing these phonetic differentiations are deep in the throat and not visible to them.

From another point of view, how people sort the world is a major inquiry in psychological sciences (Murphy, 2004). Exceptionally compelling is the order of socially and socially applicable boosts, for example, faces and voices. There is to be sure solid social strain to order gender classification precisely even in states of corrupted or not exactly finish tactile contribution as, e.g., confirm by our shame when mixing up the gender of a questioner via telephone. Luckily, such missteps are uncommon as gender classification is effectively and precisely seen through the voice alone (Whiteside, 1998), even in a nutshell on-discourse vocalizations, for example, chuckling or moans (Childers & Wu, 1991; Kreiman, 1997).

There is a significant gender dimorphism in the vocal contraction of male and female adults, influencing both the source and channel parts of voice creation (Titze, 1994). These anatomophysiological contrasts bring about various acoustical contrasts between the voices of male and female grown-up speakers and specifically the mean essential recurrence of phonation and formant frequencies (Childers & Wu, 1991). The crucial recurrence (identified with the apparent pitch) is a variable of sounds that can be effectively distinguished. When all is said in done, the basic recurrence of a sound is conversely corresponding to the size of the source, that is, children

will in gender have voices with a low recurrence or low pitch, and grown-up females will in general have voices with a high F₀ or high pitch. Nonetheless, this basic relationship doesn't generally hold. For example, Rendall, Kollias, Ney, and Lloyd (2005) demonstrated that in spite of the fact that men, by and large, have a bigger body-size and speak with a softer tone recurrence and formant frequencies than females, recurrence and subjects' gender classification can't be anticipated from body-size.

Commonly, adults can easily and precisely extract gender from acoustical data in voices. Specifically, the impression of voice gender essentially depends on the crucial recurrence (Lavner, Gath, & Rosenhouse, 2000; Mullenix, Johnson, TopcuDurgun, & Farnsworth, 1995) that is by and large higher by an octave in female than male voices; yet, pitch covers extensively among male and female voices. Utilizing social adjustment standards, in any case, Schweinberger, et al. (2008) set up that the portrayal of voice gender was moderately free of low-level acoustic data, as eventual outcomes were nullified with sinusoidal tones coordinated for principal recurrence (Schweinberger et al. 2008). This exhibited, in spite of the fact that voice pitch and gender classification are connected, other data is utilized to perceive a person's gender from his/her voice. Other acoustic parameters that may add to gender classification recognizable proof incorporate formant frequencies that mirror the separating activity of the vocal tract on voice creation (Ghazanfar & Rendall 2008; Latinus & Belin 2011) and which are likewise turn down the volume (Whiteside 1998). Studies have exhibited that gender acknowledgment likewise depends on fleeting data (Fu et al. 2004), as fleeting reversal of voices diminishes gender classification acknowledgment. These investigations exhibit that despite the fact that voice gender is by all accounts essentially perceived through voice pitch, different components can be utilized to precisely perceive gender classification, demonstrating that the impression of pitch and gender can be separated (Lattner, Meyer, & Friederici, 2005).

Functional magnetic resonance imaging (fMRI) contemplates have featured areas situated along the unrivaled fleeting sulcus (STS) answerable for preparing of voices, for both etymological and extra-semantic data in people (Belin et al. 2000) and macaques (Petkov et al. 2008). The preparing of extra-etymological parts of voices connected principally the front STS—the fleeting shaft—of the correct side of the equator, as just this locale segregated vocal from non-vocal sounds without discourse data (Belin, Zatorre, & Ahad, 2002). Examination of gender

handling of voices with fMRI has created conflicting outcomes demonstrated that female voices delivered more grounded reciprocal reaction than male voices, with a correct side of the equator predominance, particularly in the unrivaled worldly gyrus (STG), while Sokhi, Hunter, Wilkinson, and Woodruff (2005) detailed that female voice preparing included the STG while male voices delivered a bigger reaction in the privilege precuneus. Lattner, Meyer, and Friederici (2005) likewise examined pitch observation paying little respect to voice gender classification, and indicated that voice pitch included a system of districts limited shut to Heschl's gyrus. They indicated that high-pitch voices actuated a neural system like female voices while low-pitch voices instigated a bigger action in the left front cingulate gyrus; pitch judgment included the privilege prefrontal cortex.

The time course of neural voice preparing of voice attributes isn't surely known, and the writing, once more, reports conflicting outcomes. An event-related potential (ERP) study featured a voice-explicit response (VSR) 320 guys after upgrade beginning (Levy et al. 2001); the VSR, a frontal positive avoidance bigger for vocal than non-vocal improvements, was seen as consideration needy as the distinction among vocal and non-vocal sounds vanished when sound-related upgrades were unattended (Levy, Granot, & Bentin, 2003). Since that first examination, others have detailed a previous mark of voice handling (Rogier et al. 2010). While Murray et al. (2006) revealed early impacts in a living/non-living order, Charest et al. (2009) utilizing a scope of vocal improvements, including non-discourse vocalization and creature vocalization (feathered creature cries), demonstrated a particular reaction to voices beginning at 120 guys after upgrade beginning and cresting at 200 guys, for example in the idleness scope of the sound-related P2 part. Adjustments of the sound-related complex at the P2 inactivity have been depicted utilizing complex improvements: it was balanced by discourse (Tiitinen, Sivonen, Alku, Virtanen, & Naatanen, 1999) and by personality preparing in a voice acknowledgment worldview (Schweinberger, 2001). Concentrates that examined the time course of voice gender classification handling are meager, with one examination indicating a balance of the N1/P2 complex following adjustment with gender consistent vocal connectors (Zaske, Schweinberger, Kaufmann, & Kawahara, 2009). The adequacy of the N1 was diminished for male voices following adjustment to male voices, while the P2 to female voices was decreased after adjustment with female voices. Subsequently, there is little data on gender classification separation of voices, and none that have decided the spatial-worldly cerebrum designs that record this basic human ability. In the

present investigations, we investigated the time course of voice gender handling utilizing ERPs. We proposed the accompanying theories: (1) the impression of pitch and gender classification are connected yet can be separated (Lattner, Meyer, & Friederici 2005); (2) pitch handling happens sooner than gender discernment as recommended by studies showing balance of early sound-related ERPs (P50) by sound frequency (Liegeois-Chauvel et al. 1994); (3) neural action touchy to gender classification would be viewed as more noteworthy actuation to female voices, over right front destinations (Lattner, Meyer, & Friederici 2005; Zaske, Schweinberger, Kaufmann, & Kawahara, 2009). To address these theories, we quantified the neural movement identified with gender classification arrangement of voices and the job of contribute gender separation utilizing ERPs. Members performed gender order on sound clasps of voices. The job of crucial recurrence, saw as pitch, in gender classification of voices was resolved in a subsequent report, utilizing low-and piercing voices just as expected voices. To maintain a strategic distance from repetition with the covering issues in the two investigations, we present the techniques and consequences of the two examinations pursued by one general discussion.

2. METHOD

2.1 Participants

The participants of this study were 26 male (n=23) and female (n=23) advance EFL learners who were selected from a private English language institute. In fact, two intact classes were selected. Their age range of 21-26 years old. None of the participants reported any hearing problems. They all gave informed written consent and the study was approved by the institute faculties.

2.2. Auditory Stimuli

Forty English words were spoken by the participants. The words were monosyllabic. The participants voices were recorded using a Voice Recorder. The participants also spoke the words utilizing high- and low-pitched voices; speakers were trained to speak the words, making their natural pitch a higher or lower frequency, but not forcing their voices—keeping them as natural-sounding as possible, while making clearly audible changes in the pitch. All speakers were able to do so. All of the subjects completed the experiment with normal voices first and then the study with the pitch-altered voices.

2.3. Sound Analysis

With the help of professional *statistician*, *Praat software* (Boersma & Weenick 2001) was run and parameters like mean pitch, range (difference between the minimum and the maximum of pitch for each

gender) and formant frequencies (F1 to F4) plus sound duration and words' start time were measured. Moreover, two repeated measures ANOVAs were run: the first on the normal voices only, and the second one on all six categories. Voice gender was a between-subject factor, while word was a repeated factor (14 levels), when all six categories were included pitch was also a repeated factor with 3 levels.

2.4 Tasks and Design

Boosts were introduced binaurally by means of earphones at ordinary talking levels (68 ± 5 dB); between upgrade interims changed haphazardly somewhere in the range of 1,480 and 1730. The introduction request of upgrades was randomized crosswise over members. During the undertakings, a focal obsession cross was appeared on a screen 80 cm before the subjects, who were approached to keep up focal obsession and cease from making eye developments. Members squeezed one key for male voices and another for female voices (balanced subjects); in the two analyses, members were told to react as precisely and as fast as could be expected under the circumstances. Directions for the undertaking in the subsequent investigation educated the subjects that the pitch of the voices might be adjusted and, therefore may not be a legitimate prompt to segregate gender.

3. RESULTS

Comparing the normal voices, an effect was observed on the fundamental frequency (mean frequency ($p = 0.000$), which was higher for female than male voices and on the Frequency range ($P = 0.051$), which was significantly larger for male voices (Tables 1 and 2). Words affected mean frequency of the first three formants ($P = 0.041$; $P = 0.021$; $P = 0.018$ for F1, F2, and F3, respectively), in line with previous reports (Hillenbrand, Getty, Clark, & Wheeler, 1995). All other acoustical parameters, F4 frequency, sound duration and word start time were not affected by words or speakers' gender.

Table 1: Sound analysis for female participants

Female Voices				
	High-pitched	Normal	Low-pitched	
Sound duration	319.98 ± 15.9	316.19 ± 11.9	336.88 ± 5.9	
Start time	8.21 ± 2.11	11.22 ± 2.96	8.31 ± 3.16	
frequency	419.12 ± 14.9	301.19 ± 9.14	208.16 ± 3.96	
range	86.14 ± 13.98	93.19 ± 21.12	31.09 ± 4.97	

F1	711 ± 53.12	689 ± 42.12	586 ± 30.19	
F2	1819 ± 81.3	1918 ± 88.09	1799 ± 101.09	
F3	3011 ± 49.08	2987 ± 48.11	2951 ± 59.89	
F4	4088 ± 41.18	4129 ± 43.09	4098 ± 40.19	

Table 2: Sound analysis for male participants

Male Voices				
	High-pitched	Normal	Low-pitched	
Sound duration	299.99 ± 15.91	294.26 ± 12.11	304.19 ± 5.7	
Start time	10.98 ± 2.99	15.19 ± 2.98	956 ± 3.46	
frequency	389.10 ± 8.17	359.23 ± 9.51	288.06 ± 11.09	
range	94.26 ± 9.80	136.19 ± 23.69	99.69 ± 25.19	
F1	600 ± 29.49	711 ± 38.89	651 ± 64.76	
F2	1689 ± 79.08	1882 ± 72.08	1906 ± 81.13	
F3	2871 ± 47.09	2991 ± 46.11	2899 ± 55.01	
F4	4001 ± 63.19	3980 ± 51.87	3824 ± 54.64	

Moreover, analyses of frequency indicated an expected effect of gender ($P = 0.018$) and pitch ($P = 0.001$), with no interactions; female voices were on average higher pitched than male voices, and frequency was highest for high-pitch voices, while it was the lowest for low-pitch voices (Tables 1 and 2). Frequency range was still larger for male than female voices as shown by a speakers' gender effect $P = 0.037$), yet it was not modulated by pitch. Formant analysis revealed that F1, F2, and F3 frequencies differed with words. All other acoustical parameters, F4 frequency, sound duration and word start time were not affected by words, speakers' gender or pitch.

4. DISCUSSION AND CONCLUSION

Generally speaking, the researchers reported the neural correlates of voice gender perception. At first, the researchers listened to female and male voices, while performing a gender categorization; then, pitch-altered voice stimuli were included to dissociate pitch processing from higher-level gender representation processing. These two views revealed significant differences between the processing of female and male voices, both behaviorally and neurophysiologically.

At the social level, gender contrasts were found in precision just for the pitch-changed voices with pitch adjustments affecting female than male order. In the primary investigation, RTs were longer for female than male voices. In study 2, members ordered same-gender classification voices quicker than inverse gender voices, as revealed with faces (Yamaguchi, Hirukawa, & Kanazawa, 1995; Cellerino, Borghetti, & Sartucci, 2004). Additionally, members sorted the voices with the most run of the mill pitch quickest, i.e., high-pitch for female voices and low-pitch for male voices. These outcomes indicated that an individual's gender classification is to some extent gotten from principal recurrence (pitch), and give social proof that high f_0 are normal of a ladylike voice and, the other way around (Whiteside 1998; Murry & Singh 1980; Mullennix, Johnson, Topcu-Durgun, & Farnsworth, 1995). However different parameters, for example, transient data or formant recurrence, are unmistakably used to perform voice gender classification order as it stays solid even without a standard f_0 (Schweinberger, 2001; Fu, Chinchilla, Galvin, 2004). The f_0 territory could be a prompt for gender classification that remaining parts even after the pitch adjustment, as it was bigger for male voices paying little mind to pitch. Our outcomes likewise showed contrasts in the handling of female and male voices, as pitch adjustment appears to disturb female more than male voice classification. This could appear inconsistent with results indicating that that female voice observation depends more on worldly data than does the acknowledgment of male voices (Murry & Singh, 1980), yet, the improvements utilized in our two examinations were short, and transient data may have been decreased.

Examinations of the spatial–fleeting cerebrum examples were basic in uncovering contrasts in the handling of male and female voices; while investigation 1 uncovered contrasts at a few spatial–transient groups, study 2 explained that lone the impacts inside the P2 inactivity range were explicitly identified with gender classification recognition, as per Zaska et al. (2009). Early ERP contrasts, beginning at 30 ms post upgrade beginning in study 2 and at 87 ms in study 1, were owing to pitch handling, however not gender preparing as such. It has recently been exhibited that the Pa or P50, a positive potential happening in this idleness go, was touchy to upgrade recurrence and its geography reflects changes in dipole direction with expanding recurrence (Pantev et al. 1995). This adjustment in geology has been proposed to mirror the tonotopy of the essential sound-related cortex (Pantev et al. 1995). Along these lines, in the present investigations, early impacts for the most part found in geology changes because of pitch, likely reflect recurrence handling contrasts in the sound-

related cortex among high and typical to-low pitch voices. In study 1, male voices evoked a bigger N1 than female voices and this regulation by voice gender classification was additionally clear in geographical contrasts. In study 2, nonetheless, N1 was not influenced by pitch or gender, nor was any spatial distinction seen inside this inertness go. N1 mirrors the preparing of physical and worldly parts of sound-related improvements (Naatanen & Picton 1987) including recurrence (Zaska et al. 2009). Abundancy of the sound-related N1 has been demonstrated to be touchy to the physical likeness between improvements in adjustment structures (Zaska et al. 2009). N1 idleness and abundancy decline with expanding recurrence utilizing unadulterated tone upgrades, particularly for unattended tones (Crottaz-Herbette & Ragot 2000) steady with the aftereffects of our investigation 1. It has likewise been indicated that particular consideration impacts the N1 segment (Neelon, Williams, and Garell, 2006), and that consideration regarding pitch veils the N1 regulation by recurrence. This recommends the littler N1 for female voices found in study 1 relates to programmed pitch handling; this was not seen in study 2 because of consideration being coordinated away from pitch as it was not prescient and subjects were educated that pitch had been adjusted. This contrast between the two investigations is reliable with great examinations indicating that the sound-related N1 is delicate to consideration impacts (Naatanen & Picton 1987).

Female voices evoked a before and additionally bigger P2 than male voices in the two investigations: somewhere in the range of 170 and 230 ms contrasts were seen over fronto-focal mind territories that include the P2 segment. A prior P2 to female voices was accounted for in a past report (Zaska et al. 2009) and was proposed to reflect higher major frequencies in female voices. Our outcomes are in logical inconsistency with this theory, as shrill voices, with the most noteworthy crucial recurrence, evoked the most recent P2. However, Zaska, et al. (2009) additionally recommended that P2 could record a perceptual as opposed to a physical handling of female gender. In our investigations, examination between mind geographies to high-pitch male voices and typical female voices demonstrated that male voices, even with a higher f_0 , yielded littler reactions than female voices at a similar area and idleness go. Along these lines, the consolidated aftereffects of the two investigations propose that neural instruments that underlie P2 are engaged with a voice gender portrayal generally preoccupied from low-level, for example pitch, data. The P2 part has likewise been appeared to list voice handling (Charest et al. 2009; Rogier et al. 2010), as its plentifulness was higher for vocal than non-vocal sounds. Lattner, Meyer, and Friederici

(2005) showed that an infringement of audience members' desires prompted a voice-explicit mind reaction 200 ms after boost beginning. It has been proposed that P2 is a list for discourse handling, as P2 is bigger to vowels than tones (Tiitinen, Sivonen, Alku, Virtanen, & Naatanen, 1999) and is delicate to voice preparing in a voice acknowledgment worldview (Schweinberger 2001). These impacts on P2 abundance may reflect voice affectability as opposed to discourse handling. In spite of the fact that the point of our article was not to think about neural connects of voice discovery, our outcomes bolster the theory that P2 may reflect gender voice handling (Charest et al. 2009). Sokhi et al. (2005) detailed that female voices enacted the correct front STG though male voices actuated the precuneus. This was not obvious in our examination as geographies to male and female voices were similar, proposing that a typical cerebrum source is at the birthplace of the P2 for male and female voices. In any case, as fMRI information doesn't give fleeting data, the mind zones depicted by Sokhi et al. (2005) likely could be actuated at various latencies with the end goal that initiation of the STG around 200 ms prompted a bigger P2 for female voices and the precuneus enactment may happen later and drive contrasts we saw at the VSR inactivity in left back areas.

Despite the fact that the specialists discovered huge impacts in these two investigations, we recognize a few impediments. To start with, it is important that lone three voices for every gender classification were utilized in the examination, with fourteen things for every voice. This is a low number of voices, despite the fact that normal for this sort of research (for instance, five speakers for each gender classification in Zaské et al. 2009 and four speakers in Schweinberger et al. 2008). Future examinations ought to incorporate more chronicles of various voices, yet an intriguing inquiry is use voices over the age range to decide whether the age of the speaker impacts the separation of the gender classification of the speaker. Second, it is smarter to have more preliminaries per normal, to acquire significantly more clear segregation of the spatial-transient example. The danger of this would be habituation of the reactions. The way that we discovered critical impacts, with a respectable number of subjects and utilizing vigorous insights, reliable with and developing different examinations in the writing, gives us certainty that the discoveries are veridical.

Taking everything into account, this examination uncovered that sound-related ERPs record both pitch and gender preparing of voices: pitch handling begins early and is tweaked by consideration, while gender

classification separation happens around 200 ms and is likely connected with different parts of voice handling (Charest et al. 2009; Zaské et al. 2009). Therefore, we suggest that gender handling of voices has two phases. An early tonotopically-delicate stage appraises the pitch of the approaching sound; this can be a successful gauge of voice gender. In any case, when pitch data is represented, apparently contrasts at the P2 idleness stay at front-focal areas, recommending that gender segregation of voices happens at this inactivity. We recommend that genuine voice gender preparing happens at the P2 dormancy while pitch handling, which could be an increasingly fast surrogate for gender classification preparing, happens a lot prior.

REFERENCES

- [1] Belin, P., Zatorre, R. J., Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Brain Res Cogn Brain Res*, 13(1), 17–26.
- [2] Boersma, P., Weenick, D. (2001). Praat, a system for doing phonetics by computer. *Glott Int*, 5(9/10), 341–345.
- [3] Cellerino, A., Borghetti, D., & Sartucci, F. (2004) Sex differences in face gender recognition in humans. *Brain Res Bull*, 63(6), 443–449.
- [4] Charest, I., Pernet, C.R., Rousselet, G.A., Quinones, I., Latinus, M., Fillion-Bilodeau, S., Chartrand, J.P., & Belin, P. (2009). *Electrophysiological evidence for an early processing of human voices*. BMC: Neurosci
- [5] Childers, G., and Wu, K. (1991). Gender recognition from speech. Part II: fine analysis. *J. Acoust. Soc. Am*, 90, 1841–1856.
- [6] Crottaz-Herbette, S., & Ragot, R. (2000). Perception of complex sounds: N1 latency codes pitch and topography codes spectra. *Clin Neurophysiol*, 111(10), 1759–1766.
- [7] Fu, Q. J., Chinchilla, S., Galvin, J.J. (2004). The role of spectral and temporal cues in voice gender discrimination by normal-hearing listeners and cochlear implant users. *J Assoc Res Otolaryngo*, 15(3), 253–260.
- [8] Ghazanfar, A. A., & Rendall, D. (2008). Evolution of human vocal production. *Curr Biol*, 18(11), R457–R460.
- [9] Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of

American English vowels. *J Acoust Soc Am*, 97, 3099–3111.

[10] Kreiman, J. (1997). “Listening to voices: theory and practice in voice perception research”, in *Talker Variability in Speech Processing*, eds K. Johnson, and J. W. Mullennix (San Francisco: Morgan Kaufmann Publishers), 85–108.

[11] Latinus M., & Belin, P. (2011). Human voice perception. *Curr Biol* 21(4), R143–R145.

[12] Lattner, S., Meyer, M.E., Friederici, A. D. (2005). Voice perception: sex, pitch, and the right hemisphere. *Hum Brain Mapp* 24(1), 11–20.

[13] Lavner, Y., Gath, I., and Rosenhouse, J. (2000). The effects of acoustic modifications on the identification of familiar voices speaking isolated vowels. *Speech Commun*, 30,9–26.

[14] Levy, D. A., Granot, R., Bentin, S. (2001) Processing specificity for human voice stimuli: electrophysiological evidence. *Neuroreport*, 12(12), 2653–2657.

[15] Liegeois-Chauvel, C., Musolino, A., Badier, J.M., Marquis, P., & Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. *Electroencephalogr Clin Neurophysiol*, 92(3), 204–214.

[16] Mullennix, J. W., Johnson, K. A., TopcuDurgun, M., & Farnsworth, L. M. (1995). The perceptual representation of voice gender. *J. Acoust. Soc. Am*, 98, 3080–3095.

[17] Mullennix, J. W., Johnson, K.A., Topcu-Durgun, M., Farnsworth, L.M. (1995). The perceptual representation of voice gender. *J Acoust Soc Am* 98(6):3080–3095.

[18] Murphy, G. L. (2004). *The Big Book of Concepts*. Cambridge: MIT Press

[19] Murray, M.M., Camen, C., Gonzalez Andino, S.L., Bovet, P., & Clarke, S. (2006). Rapid brain discrimination of sounds of objects. *J Neurosci*, 26, 1293–1302.

[20] Murry, T., Singh, S. (1980). Multidimensional analysis of male and female voices. *J Acoust Soc Am*, 68(5), 1294–1300.

[21] Naatanen, R., & Picton, T. (1987), The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, 24(4), 375–425.

[22] Neelon, M. F., Williams, J., & Garell, P. C. (2006). The effects of auditory attention measured from human electrocorticograms. *Clin Neurophysiol*, 117(3), 504–521.

[23] Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., & Elbert, T. (1995). Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *Electroencephalogr Clin Neurophysiol*, 94(1), 26–40.

[24] Rendall, D., Kollias, S., Ney, C., & Lloyd, P. (2005). Pitch (f0) and formant profiles of human vowels and vowel-like baboon grunts: the role of vocalizer body size and voice acoustic allometry. *J. Acoust. Soc. Am*, 117, 944–955.

[25] Rogier, O., Roux, S., Belin, P, Bonnet-Brilhault, F., & Bruneau, N. (2010). An electrophysiological correlate of voice processing in 4- to 5-year-old children. *Int J Psychophysiol*, 75(1), 44–47.

[26] Schweinberger, S. R., Casper, C., Hauthal, N., Kaufmann, J. M., Kawahara, H., Kloth, N., Robertson, D.M. C., Simpson, A. P., and Zaske, R. (2008). Auditory adaptation in voice perception. *Curr. Biol*, 18, 684–688.

[27] Schweinberger, S.R. (2001). Human brain potential correlates of voice priming and voice recognition. *Neuropsychologia* 39(9), 921–936.

[28] Schweinberger, S.R., Casper, C., Hauthal, N., Kaufmann, J.M., Kawahara, H., Kloth, N., Robertson, D.M., Simpson, A.P., Zaske, R. (2008). Auditory adaptation in voice perception. *Curr Biol*, 18(9), 684–688

[29] Sokhi, D. S., Hunter, M.D., Wilkinson, I.D., Woodruff, P. W. (2005). Male and female voices activate distinct regions in the male brain. *Neuroimage*, 27(3), 572–578.

[30] Tiitinen, H., Sivonen, P., Alku, P., Virtanen, J., & Naatanen, R. (1999). Electromagnetic recordings reveal latency differences in speech and tone processing in humans. *Brain Res Cogn Brain Res*, 8(3), 355–363.

[31] Titze, I. R. (1994). *Principles of Voice Production*. Englewood Cliffs: PrenticeHall.

[32] Whiteside, S. P. (1998). Identification of a speaker’s sex: a study of vowels. *Percept. Mot. Skills*, 86, 579–584.

[33] Yamaguchi, M. K., Hirukawa, T., & Kanazawa, S. (1995). Judgment of gender through facial parts. *Perception* 24(5), 563–575.

[34] Zaske, R., Schweinberger, S.R., Kaufmann, J.M., & Kawahara, H. (2009). In the ear of the beholder: neural correlates of adaptation to voice gender. *Eur J Neurosci* 30(3), 527–534.