A Systematic Analysis of Speech Perception and Sex Differences

Ashley Kim¹, Lyndsey Kim²
¹Union County Vocational-Technical High School, New Jersey, United States
²The Academy for Allied Health Sciences, New Jersey, United States

Abstract
Speech perception is the process by which the sounds of language are heard, recognized, interpreted, and understood. Because understanding spoken language is an integral part of our social interactions, communications, and learning, speech perception has been an active research topic in the fields of psychology, cognitive science, and linguistics. Research in speech perception seeks to understand how humans recognize speech sounds and use this information to understand spoken language. In this review paper, we first overview some of the brain areas underlying speech perception. We then discuss previous findings that have found significant sex differences in speech perception and those that rather suggested similarities in speech perception between women and men. Reviewing the discrepancy between the research findings will allow us to determine the degree to which sex differences impact the way individuals perceive and understand speech sounds. Although speech perception is mainly based on auditory processing, vision often plays an important role as well. In this review paper, we also include previous studies on such multimodal processing in speech perception, because, in everyday conversations in which we are routinely engaged, we often benefit from using different sensory information from different modalities (e.g., audition and vision). We conclude that while there are interesting findings on sex differences in speech perception, particularly in females, more systematic and quantitative evaluation should also be made, before making any clear conclusions about sex differences in speech perception in general.

Keywords: Auditory Processing; Cognitive Science; Literature Review; Speech Perception

1. Introduction

You're waiting for a bus on a busy, crowded street. All around you are different sounds of trains, cars, traffic signals, and music coming from store fronts. Among all these, one of the sounds you suddenly realize that you are hearing is someone nearby speaking on the phone, and that you are listening to what the person says in a different way from how you hear all the other sounds from different sources that have been there for quite a while. Researchers show that when we hear sounds of understandable language, our brains go through a highly specialized, automatic process to pick up on the spoken language. Our brains also seem to react differently than they do when we hear non-speech sounds or people talking in languages we do not understand. When we hear someone talking in a familiar language, our brain quickly shifts to pay attention to the speech sounds that are heard, by turning them into small meaningful elements for speech perception, words, and sentences, then processing the meanings of what is being said.
Speech perception is the process by which the sounds of language are heard, recognized, interpreted, and understood. Because understanding spoken language is an integral part of our social interactions, communications, and learning, speech perception has been an active research topic in the fields of psychology, cognitive science, and linguistics. Research in speech perception seeks to understand how humans recognize speech sounds and use this information to understand spoken language. Typically, the process of speech perception begins from the brain’s auditory system that allows us to hear sounds. Unlike other sound signals, a subset of sounds that belong to the structure of language a listener has acquired is then carried to specialized processing streams for further processing to extract acoustic cues and phonetic information, followed by higher-level language processes such as word recognition or semantic processing.

As we always experience how much each of us differs in the way we perceive and understand spoken language due to various reasons, dozens of scientific research studies have been conducted in the past few decades, to understand the foundations of individual differences in speech perception. Particularly, many previous studies have explored sex differences in speech perception, in terms of female and male observers’ ability to understand spoken sound and underlying brain functions and structure, both in the clinical population that shows impaired speech perception ability and typically developing population. In many of these studies, sex differences in speech perception have been abundantly reported. For example, it is widely suggested that throughout the course of development and into adulthood, female observers are generally better than male observers at perceiving and understanding speech sounds (Johnson et al., 1988; Dancer et al., 1994; Watson et al., 1996; Strelnikov et al., 2009). In addition, neuroanatomical studies have indicated that when presented with visual speech, females have stronger activation in brain areas associated with speech perception than males (Ruytjens et al., 2006, 2007). Neuroanatomical studies have also suggested gender differences in lateralization of speech processing (e.g., Shaywitz et al., 1995; Jaeger et al., 1998), where females have more bilateral processing for word recognition (e.g., Walla et al., 2001) and for tasks involving phonology and syntax (Pugh et al., 1996; Jaeger et al., 1998). However, sex differences in speech perception still remain controversial because results are quite mixed and have not been systematically reviewed or assessed. For example, there are also considerable research findings that suggest an absence of sex differences in perceiver’s speech perception ability and neural activities (e.g., Baxter et al., 2003; Clements et al., 2006; Frost et al., 1999; Hund-Georgiadis et al., 2002; Sommer et al., 2004). Furthermore, the sources of such sex-related differences have still not well understood. The current paper reviewed the previous research studies that examined sex differences in behavioral abilities or brain structure and functions, with the aim of determining the degree of agreement among these studies on speech perception. Identifying sex-related differences in behavioral and neural processing underlying the speech perception will provide useful information for developing more precise models and theories on speech perception, which will help researchers to make more accurate predictions for speech learning or therapy and building better computer systems for speech recognition, translation, or hearing aids.

In this review paper, we first overview some of the brain areas underlying speech perception. We then discuss previous findings that have found significant sex differences in speech perception and those that rather suggested similarities in speech perception between women and men. Reviewing the discrepancy between the research findings will allow us to determine the degree to which sex differences impact the way individuals perceive and understand speech.
sounds. Although speech perception is mainly based on auditory processing, vision often plays an important role as well. For example, it has been shown that lip-reading using visual information we receive by reading each other’s speech and interpreting the movements of the lips and adjacent regions in the face affects how we perceive speech sounds we hear. In this review paper, we also include previous studies on such multimodal processing in speech perception, because, in everyday conversations in which we are routinely engaged, we often benefit from using different sensory information from different modalities (e.g., audition and vision).

2. Methods
2.2 Search strategy
A comprehensive electronic literature search was performed using PubMed in July 2021. The search identified behavioral and neuroscience research studies investigating sex differences in speech perception, published between the years 1993 and 2021. The key search terms for the title and/or the abstract of journal articles included sex difference(s) or gender difference(s), and speech perception. The title and abstract of all articles shown in the search results were screened and assessed thoroughly. To discuss sex differences in the multimodal processing of speech perception, additional references reviewed in this paper were further retrieved from some of the selected articles shown in the search results.

2.2.2 Selection criteria
From the Pubmed search results, only English publications from peer-reviewed journals were selected. Twenty-three papers were identified from the search results. In this review paper, we focus on healthy adults. Studies involving children as participants (Bednarek D, Saldaña D, & García, 2009; Norrelgen, Lacerda, & Forssberg, 2001; Ross et al., 2015) were removed from the review. For the 16 remaining studies, it was also ensured that direct contrast between men and women was assessed, and that speech perception was tested, instead of other cognitive abilities that are related to speech perception (e.g., go-no-go task, Kriengwatana et al., 2015; Lapenta et al., 2012).

3. Results
3.3 Brain mechanisms underlying speech perception
The hemispheric asymmetry, different patterns of functional dominance between the left and right hemispheres have been studied in many previous studies (e.g., McGettigan & Scott, 2012; Scott & McGettigan, 2013). Particularly, speech perception is traditionally known to be left-lateralized in the human brain (Knecht et al., 2000; Markus & Boland, 1992; Patel, & Wise, 2008; Scott et al., 2000; Wise, 2003), which means that the left hemisphere of the brain is specialized in perceiving speech sounds.

A number of neuroimaging studies (Jacquemot et al., 2003) have investigated the brain regions supporting speech perception (Figure 1). Perceiving consonants and syllables was observed to engage the brain activation in the posterior superior temporal sulcus (STS) and the superior temporal gyrus (STG) and the middle temporal gyrus (MTG), with some suggestion of a greater level of activation in the left hemisphere compared to the right hemisphere (Obleser et al., 2007), consistent with the notion of the left lateralization of speech perception. Moreover, the supramarginal gyrus and the inferior frontal gyrus (IFG) are also known to be associated with
speech perception (Schirmer et al., 2004). Thus, perceiving speech sound (e.g., phonemes) appears to be supported by the function of brain areas, including the STS/STG, MTG, IFG, and supramarginal gyrus.

Unlike the processing speech sounds that mostly involve auditory processing, processing single words involve semantic representations to understand the meaning of what is heard. Neuroimaging studies of speech processing in the level of understanding words, temporal lobe activation both in the left and right hemispheres has been reported (Binder et al., 2000; Price et al., 1996, Price et al., 1992), which included the STS and MTG regions that showed activations when people heard spoken words (Eckert et al., 2008, Bozic et al., 2010, Fonteneau et al., 2015, Sohoglu et al., 2012, Specht & Reul, 2003, Burton & Small, 2006). In processing spoken words, there is converging evidence that both left, and right temporal cortices support single word processing, which is inconsistent with the notion of left lateralization for speech perception in general. For example, one piece of convincing evidence comes from the previous study in which patients had the left hemisphere anesthetized but still could understand meanings of spoken words with only their right hemisphere active (Hickok et al., 2008). Therefore, it appears that the degree to which the brain is lateralized for different levels of speech perception varies, and how the hemispheric lateralization depends on the type of speech processing will need to be examined by quantitatively comparing the involvement of the left and right hemispheres in future studies. For further semantic processing that provides listeners with the meaning of speech sounds, including words and sentences, neuroimaging studies have suggested important roles of the angular gyrus (Binder et al., 2009; Price et al., 2015, Graves et al., 2010; Lee et al., 2007) and the left IFG (Thompson-Schill et al., 1997; Whitney et al., 2011).

3.3.3 Sex differences in speech perception
A set of neuroscience studies have suggested sex-related differences in the pattern of brain functions during speech processing (e.g., Shaywitz et al., 1995; Jaeger et al., 1998). For example, females were observed to have more bilateral processing for tasks that involved speech sound perception (Pugh et al., 1996; Jaeger et al., 1998) or word recognition (e.g., Walla et al., 2001). Moreover, it was observed that females and males showed differences in the brain structures that are related to the gray matter volumes and tissue density, mostly in the brain regions, including the speech and language areas (Ruigrock et al., 2014). Different connectivity patterns between the left and right hemispheres were also observed, with men showing greater

![Figure 1. Brain areas that mediate speech perception.](image-url)
Communications between the hemispheres compared to women (Ingalhalikar et al., 2014). Finally, it has also been shown that brain activation that is evoked by speech sounds is initiated earlier in females’ brains than males’ brains (Näätänen & Picton, 1987; Woods, 1995), suggesting that female brains tend to respond to speech stimuli earlier than male brains and sex-differences in speech perception may exist from very early stages of auditory and speech processing.

Speechreading refers to the ability to recognize and distinguish speech sounds based on visual cues available from the lip and facial movements. Speechreading is an important skill that facilitates our speech perception, especially when the auditory cue itself is not sufficient (e.g., noisy environments or impaired hearing, etc.). In general, previous research suggests that females are better speech-readers than males (e.g., Johnson et al., 1988; Dancer et al., 1994; Watson et al., 1996; Strelnikov et al., 2009), as well as females tend to be more actively engaged in speechreading than males (Berndl et al., 1986; Johnson et al., 1988).

Not only better ability to read speech from lip movement, but females have also been reported to be more influenced by the visual signal in audio-visual speech perception (Aloufy et al., 1996; Öhrström & Traunmüller, 2004; Irwin et al., 2006). When perceiving visual speech, females also show greater activation of brain areas that are associated with speech perception, compared to males (Ruytjens et al., 2006, 2007). Integration of visual and auditory cues about speech has been extensively studied by using the phenomenon, called McGurk stimuli. A set of stimuli for the McGurk effect are created by combining non-matching visual speech information and auditory speech information. The McGurk effect is an audiovisual illusion in which the presentation of incongruent visual inputs (lip movements) can greatly modify the speech sound that is heard (McGurk & MacDonald, 1976; Saint-Amour et al., 2007). As illustrated in Figure 2, when an observer hears a person say “Ba,” the speech sound that the observer perceives can be dramatically changed by the visual information that is provided at the same time. For example, when the person’s lip movement also shows “Ba,” then audio and visual information match, providing the perception of “Ba” in the observer’s speech processing (Figure 2A). However, when the person’s lip movement is manipulated to show “Fa,” then the observer tends to perceive the original “Ba” sound as “Fa” sound (Figure 2B). The McGurk effect provides compelling evidence that visual information dominates auditory information in some speech perception. Such audiovisual integration in speech perception has been examined to understand any systematic differences between female and male observers. Some researchers have reported a stronger McGurk effect in females than males (Aloufy et al., 1996; Öhrström & Traunmüller, 2004). Irwin et al. (2006) showed that women displayed the greater
influence of visual speech on heard speech, even when there were no clear speech-reading abilities between women and men.

In contrast to these studies that have reported sex differences in brain functions and structure related to speech perception, however, the existence of sex differences for speech perception remains controversial, because considerable research has also shown an absence of sex differences in both behavioral performance (e.g., Baxter et al., 2003; Clements et al., 2006) and neuroanatomical measurements (e.g., Frost et al., 1999; Hund-Georgiadis et al., 2002; Sommer et al., 2004). These studies suggest that there are much more neurobiological and behavioral similarities than differences between men and women during speech and language processing.

Potential factors for such discrepancy can be associated with task types or stimuli that have been used in different studies, as well as perceivers’ biological cycles. For example, some researchers have shown that young females are better speech-readers in general than young adult males for word processing, but not for sentence processing (Dancer et al., 1994; Watson et al., 1996) or processing and recognizing meaningless syllables (Strelnikov et al., 2009). Furthermore, one study (Wadnerkar et al., 2008) reported that sex differences in the ability to listen to speech sounds were found from the comparison made between males and females when the female perceivers were in a particular menstrual cycle (e.g., low estrogen and progesterone), but not the other (e.g., high estrogen and progesterone). This result suggests that hormones can play important roles in influencing the way the brain processes speech sounds in women. Thus, such studies also highlight the importance of considering various biological factors that differentially affect speech processing in females and males when evaluating sex differences. Together, these studies may provide some insights into why studies on sex differences in speech perception are indecisive.
4. Conclusion
To conclude, there are interesting findings on sex differences related to speech perception. To summarize, females tend to show better abilities and greater sensitivity to speech stimuli, better abilities to read speech using visual cues, stronger influences by visual cues on auditory perception of speech sounds, when visual and auditory cues do not match each other. Together, sex appears to be an important biological factor that needs to be considered when designing experiments on speech perception. For example, it would be important to have an equal number of female and male participants when designing and conducting psychological or neuroscientific research studies. Otherwise, some interesting patterns of results on speech perception in healthy adults might be weakened or undesirably biased, resulting in potentially misleading conclusions. At the same time, however, there are some studies reporting that have failed to confirm sex differences, depending on groups of participants recruited, types of specific tasks for speech perception, or kinds of stimuli used for the experiments. Therefore, more systematic and quantitative evaluation should also be made, before making any clear conclusions about sex differences in speech perception in general. There are many different biological factors that need to be controlled at the same time by researchers to evaluate sex differences in speech perception, including hormonal cycle effects or ages, and so on. Better understanding both differences and similarities between females and males would also need to be determined at the same time. This will provide us with useful information for designing different speech and language programs to teach, evaluate, or rehabilitate abilities to understand speech in different individuals and developing better automatic translators as tools for hearing aids or artificial intelligence.
References


coverage: the example of human papillomavirus vaccine following public concerns about vaccine safety.” BMC public health vol. 15 1233. 12 Dec. 2015, doi:10.1186/s12889-015-2575-7


Marshall Graves, J.A., Peichel, C.L. Are homologies in vertebrate sex determination due to shared ancestry or to limited options?. Genome Biol 11, 205 (2010). https://doi.org/10.1186/gb-2010-11-4-205


Ross, Lars A et al. “Sex differences in multisensory speech processing in both typically developing children and those on the autism spectrum.” Frontiers in neuroscience vol. 9 185. 27 May. 2015, doi:10.3389/fnins.2015.00185


