Speech Perception in Schizophrenia Patients

Authors: Cao Yi, Yang Xiaohu

Date: 2018-12-05T00:00:00+00:00

Abstract

Schizophrenia is a common psychiatric disorder characterized by multifaceted symptoms, among which language abnormalities constitute one of the core symptoms of cognitive impairment in patients with schizophrenia. This paper focuses on speech perception in patients with schizophrenia, briefly reviewing behavioral and neuroscientific experiments on speech perception conducted both domestically and internationally from segmental and suprasegmental perspectives, and points out that China should intensify research on Mandarin speech perception in Chinese patients with schizophrenia.

Full Text

Speech Perception in Schizophrenia Patients*

CAO Yi[1], YANG Xiaohu[1,2]
(1 Speech-Language-Hearing Center, School of Foreign Languages, Shanghai Jiao Tong University, Shanghai 200240, China)
(2 Institute of Social Cognitive and Behavioral Sciences, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract

Schizophrenia is a common psychiatric disorder characterized by multiple symptoms, among which language abnormalities constitute one of the core features of cognitive impairment in schizophrenia patients. This article focuses on speech perception in schizophrenia patients, briefly reviewing behavioral and neuroscientific experiments on speech perception in schizophrenia conducted both domestically and internationally from the perspectives of segmental and suprasegmental levels, and points out that China should increase research efforts into Mandarin Chinese speech perception among Chinese schizophrenia patients.

Keywords: schizophrenia; speech perception; neuroscience

Introduction

Schizophrenia is a complex, heterogeneous syndrome of behavioral and cognitive dysfunctions that profoundly affects patients, their families, and society. The symptoms of schizophrenia are generally considered to include positive symptoms, negative symptoms, and cognitive impairment (Owen, Sawa, & Mortensen, 2016). As one of the primary symptoms of schizophrenia, cognitive impairment encompasses not only neuropsychological domains such as attention, memory, processing speed, and reasoning, but also higher-level aspects of social cognition, including mental activities that guide social behavior such as attribution, intention, action, and emotion (van, Kenis, & Rutten, 2010). Language, as the foundation of human behavior, constitutes an essential component of cognitive function, and many schizophrenia patients exhibit language abnormalities (Covington et al., 2005). For patients, language barriers not only affect social communication but also reduce quality of life (Bellani, Perlini, & Brambilla, 2009). Therefore, studying language abnormalities in patients is of significant importance.

From a linguistic perspective, current domestic and international research on language abnormalities in schizophrenia patients can be broadly categorized into various linguistic levels including phonology, morphology, syntax, semantics, and pragmatics (Bellani et al., 2009; Covington et al., 2005), among which examining language abnormalities through speech perception offers a relatively direct approach. At present, scholars both domestically and internationally have focused not only on the manifestations of speech perception abnormalities in schizophrenia but also on the neural mechanisms underlying speech perception deficits, attempting to reveal the causes of speech perception abnormalities through examining brain anomalies in patients. This article provides a comprehensive review of research on speech perception in schizophrenia patients and offers perspectives on future research trends.

2. Segmental Perception in Schizophrenia Patients

Speech can be described using segmental and suprasegmental levels, where the segmental level refers to vowels and consonants, and the suprasegmental level involves speech features that extend beyond individual consonants and vowels, primarily including tone, stress, and intonation, with phonetic foundations in the dynamic patterns of pitch, duration, and intensity (Clark & Yallop, 1990; Ladefoged & Johnson, 2011). This article reviews speech perception performance in schizophrenia from both segmental and suprasegmental perspectives.

Kugler and Caudrey (1983) conducted a phoneme perception study with schizophrenia patients and healthy controls, in which participants were asked to judge whether heard syllables (such as /ba/-/ba/, /ba/-/da/) were identical, and found that patients exhibited significant perceptual impairment. Similarly, Cienfuegos, March, Shelley, and Javitt (1999) used behavioral experiments to examine phoneme perception in schizophrenia patients and healthy controls, asking participants to discriminate between natural /ba/, /da/ syllables and a synthesized continuum from /ba/ to /da/. The results showed that both patient and control groups could identify natural /ba/ and /da/ syllables, but patients performed significantly worse when identifying synthesized syllables, suggesting impaired categorical perception of speech sounds in schizophrenia.

Hugdahl et al. (2012) further revealed the relationship between auditory hallucinations and phoneme perception through dichotic listening tests. In the experiment, different syllables such as /ba/-/da/ were simultaneously presented to the left and right ears of schizophrenia patients. The results showed a significant negative correlation between the severity of patients' hallucination symptoms and correct perception rates in the right ear. While normal speech perception primarily occurs in the left hemisphere, manifesting as a right-ear advantage, patients' right-ear perception accuracy decreased as hallucination symptoms worsened. The researchers proposed that the brain region processing right-ear speech perception overlaps with that generating auditory hallucinations, suggesting that hallucinations originate from aberrant speech perception in the left hemisphere, thereby providing support for the speech perception model of auditory hallucinations.

Tsao, Chiang, and Liu (2013) used synthesized affricate and fricative syllables as speech stimuli to test schizophrenia patients with and without auditory hallucinations, along with healthy controls. They found that patients' perception of affricates was worse than that of healthy controls, and patients with hallucinations performed even worse than those without. This indicates that schizophrenia patients exhibit reduced sensitivity in speech perception and that auditory hallucinations can cause consonant perception impairment.

To further understand the physiological basis of phoneme perception abnormalities and the cortical regions involved, Kasai et al. (2002) and Kasai et al. (2003) used ERP and MEG respectively to record phoneme perception in schizophrenia patients and healthy controls. Both studies employed the same three stimulus sets designed to elicit mismatch negativity (MMN) and magnetic mismatch negativity (MMNm) responses to changes in pure tone duration, Japanese vowel /a/ duration, and vowel phoneme identity. The first study found that phoneme changes elicited the smallest MMN amplitudes in patients compared to changes in pure tone and vowel duration, and that phoneme changes produced smaller bilateral MMN amplitudes in patients relative to healthy controls. The authors suggested that patients' frontotemporal cortical networks were compromised. The second study revealed that patients exhibited smaller magnetic field strengths for all three types of MMNm responses compared to healthy individuals, with phoneme changes producing the most significant differences. Additionally, patients showed no abnormal asymmetry in MMNm power, providing physiological evidence for bilateral auditory cortex impairment. The study also identified damage to neuronal clusters in the planum temporale involved in MMNm generation.

The aforementioned studies have explored segmental perception performance in schizophrenia patients through behavioral and neuroscientific experiments, revealing impairments across various levels of segmental perception that are associated with auditory hallucinations. Additionally, damage to patients' auditory cortex may underlie segmental perception deficits. However, despite providing preliminary conclusions, several aspects of existing research warrant further scrutiny and investigation. Regarding participant selection, most experiments have involved small sample sizes; age ranges have been too broad to effectively control for effects of cognitive aging on speech perception; diagnostic criteria for schizophrenia have been inconsistent, with some studies using DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, Fourth Version) while others used ICD-10 (International Classification of Diseases, Tenth Version); illness duration has varied too widely to control for disease progression effects on perceptual abilities; medication status during experiments has been inconsistent, making it difficult to control for drug effects; most studies have used native English speakers, lacking examination of other languages; and some experiments have not assessed participants' auditory abilities, making it difficult to control for hearing loss effects on speech recognition. Although Kasai et al. (2002) and Kasai et al. (2003) excluded participants with hearing impairment, they did not specify the exclusion criteria. Regarding experimental stimuli, existing research has been limited to narrow categories of segmental features, such as syllables composed of stops and vowels, synthesized affricate and fricative consonants, and single vowels, which cannot comprehensively reflect whether patients exhibit perception abnormalities for all types of phonemes. For instance, whether patients show impairments in perceiving more fine-grained vowel categories (such as front/back, high/low vowels) and consonant categories (such as nasals, approximants) remains to be investigated.

3. Suprasegmental Perception in Schizophrenia Patients

Suprasegmental features are often associated with semantics and thus serve a phonemic distinctive function. These features manifest as the rhythmic undulation of sounds in speech flow, leading some phoneticians to refer to them as prosodic features (华维芬, 1993). Prosody can be divided into linguistic and emotional prosody (Hoekert, Kahn, Pijnenborg, & Aleman, 2007). Linguistic prosody involves semantic decision-making, such as emphasizing main parts of sentences and presenting information through statements or questions. Emotional prosody refers to emotion-laden intonation, which is crucial for perceiving emotional states and speakers' intentions. Research on linguistic prosody perception in schizophrenia patients has primarily focused on three acoustic indices: pitch, duration, and intensity. Studies on emotional prosody have mainly addressed two aspects: examining acoustic variables of various emotions and attempting to link the ability to comprehend emotional prosody with specific brain regions (Edwards, Jackson, & Pattison, 2002; Hoekert et al., 2007).

3.1 Pitch Perception

Pitch refers to the highness or lowness of sound and is the perceptual correlate of fundamental frequency (F0) (Clark & Yallop, 1990). Research on pitch perception in schizophrenia patients falls into two categories: studies of static pitch through measuring pitch perception abilities, and studies of dynamic pitch, which measure the ability to perceive pitch changes, such as those occurring in lexical tones. The following sections review research on pitch perception performance in schizophrenia from both static and dynamic perspectives.

Dondé et al. (2017) identified the tone-matching task as the primary method for assessing static pitch perception abilities in patients. This task is simple, accurate, and reliable: participants hear pairs of non-speech sounds that are either identical or different in pitch and respond by pressing "same" or "different" buttons, with researchers evaluating performance by calculating response accuracy rates. Their systematic review and meta-analysis of studies on tone-matching abilities in schizophrenia found that patients performed significantly worse than healthy controls on tone-matching tasks.

Recognizing that most research has focused on native English-speaking schizophrenia patients, Yang et al. (2012) investigated native Mandarin-speaking schizophrenia patients. The authors found that, like their English-speaking counterparts, Mandarin-speaking schizophrenia patients exhibited tone-matching deficits. Specifically, Mandarin-speaking patients required an average pitch difference of 20% to achieve the same accuracy rate that controls attained with only a 2.5% pitch difference.

Meanwhile, a series of studies have explored the neural mechanisms underlying these deficits. Javitt (2000) found that deviant auditory stimuli with 20% pitch differences elicited MMN amplitudes in patients similar to those elicited by 5% pitch differences in healthy controls, indicating that patients' poor pitch discrimination ability is associated with reduced MMN amplitudes evoked by deviant stimuli. Furthermore, EEG, MEG, and fMRI studies have demonstrated that MMN originates in the primary auditory cortex, leading to the inference that patients' primary auditory cortex is damaged. Specifically, the impairment may involve abnormal neurotransmission regulated by NMDA (N-methyl-D-aspartate) receptors in the auditory cortex. Javitt, Spencer, Thaker, Winterer, and Hajós (2008) reiterated this view, stating that MMN abnormalities can reflect primary auditory cortex damage.

While Javitt (2000) and Javitt et al. (2008) explained patients' poor pitch discrimination from the perspective of functional impairment in the primary auditory cortex, Sweet et al. (2007) offered an explanation based on structural damage. Sweet et al. (2007) found structural damage in the feedforward circuits of the primary auditory cortex in schizophrenia patients, specifically lower density of synaptophysin-immunoreactive puncta (markers of axon terminals). They analyzed that this damage could lead to abnormal activation propagation in the primary auditory cortex following auditory stimulation, thereby causing deficits in both MMN and tone-matching abilities.

Todd et al. (2008) further discovered that reduced MMN amplitudes elicited by pitch changes were significant only in schizophrenia patients with long illness duration, but not in first-episode or short-duration patients, suggesting that pitch-change-evoked MMN could serve as an index of auditory cortex pathology severity.

Research on dynamic pitch can be divided into two types. The first involves studying pitch changes in English sentences. For instance, Matsumoto et al. (2006) examined performance on three tasks—discriminating within-sentence pitch, sentence-final pitch, and sentence rhythm—in schizophrenia patients and healthy controls. They found that, compared to healthy individuals, schizophrenia patients showed no abnormalities in understanding sentence-final pitch changes, such as those in question intonation, but performed significantly worse in understanding within-sentence pitch changes.

The second type involves research on lexical tones in Mandarin Chinese. Tone is the pitch movement pattern of syllables or words (word groups), including both the contour form of pitch rises and falls (i.e., tone shape) and relative pitch level characteristics (曹剑芬, 2002), making tone essentially a complex dynamic pitch change. Regarding the role of pitch in tonal versus non-tonal languages, Wang et al. (2018) reviewed the literature and found that in tonal languages like Mandarin, the primary acoustic feature constituting tone is fundamental frequency, which can distinguish semantic meanings of identical syllables or word groups, whereas in non-tonal languages like English, pitch mainly conveys pragmatic information such as emphasis, sentence modality, or emotion. Therefore, pitch contours play important roles in speech perception for both tonal and non-tonal languages.

Tsao et al. (2013) tested schizophrenia patients with and without auditory hallucinations and healthy controls using Mandarin tones on the vowel /i/ as stimuli. The results revealed that patients' accuracy in perceiving the four Mandarin tones was significantly lower than that of healthy controls, with patients experiencing hallucinations performing even worse. Yang et al. (2012) compared schizophrenia patients and healthy controls on tone matching tests, distorted Chinese tune tests, word discrimination, and word identification tasks, finding that schizophrenia patients performed significantly worse on every test. The authors concluded that schizophrenia patients suffer severe impairments in both lexical tone and auditory word processing, with these two deficits being significantly correlated. Wang et al. (2018) designed four types of sentence stimuli: sentences with correct tone, syntax, and semantics; sentences with incorrect tone (acoustically characterized by flat F0 contours) but correct syntax and semantics; sentences with correct tone but incorrect syntax and semantics; and sentences with incorrect tone, syntax, and semantics. Participants were asked to write down the sentences they heard, and comprehension accuracy was assessed by counting correctly identified keywords. The results showed that schizophrenia patients performed significantly worse than healthy controls in understanding all four sentence types, and that more severe hallucinations were associated with poorer sentence comprehension. The authors speculated that internal noise caused by auditory hallucinations may degrade speech perception and understanding.

These three studies demonstrated tone perception impairments in patients through behavioral experiments. Wang et al. (2017) analyzed the phenomenon of impaired tone perception from a neuroscientific perspective, using an auditory oddball paradigm (with standard and deviant stimuli being pure tones at 1000 Hz and 2000 Hz respectively) to study Chinese schizophrenia patients. They found that Chinese schizophrenia patients showed smaller P300 effects compared to healthy controls, demonstrating impaired ability to discriminate pure tones of different pitches. Since P300 effects can reflect the ability to identify different pure tones across frequencies, this ability may constitute the foundation of tone awareness.

Overall, researchers have conducted several studies on schizophrenia patients' perception of the suprasegmental feature of pitch using behavioral and neuroscientific experiments, covering both static and dynamic pitch. Patients have shown poorer performance than normal individuals in perceiving both forms of pitch. Neuroscientific experiments have identified MMN and P300 as indicators that can reflect pitch perception impairments, which may result from structural and functional damage to the primary auditory cortex. Furthermore, impairments in basic pitch perception ability may lead to auditory word perception errors (Yang et al., 2012), sentence perception errors (Matsumoto et al.,. 2006; Wang et al., 2018), and deficits in higher-level cognitive abilities such as emotion perception (Dondé et al., 2017), thereby affecting language communication. For example, patients' pitch perception impairments correlate with poor performance on auditory word identification and discrimination tasks (Yang et al., 2012), patients have difficulty understanding sentence emphasis (Matsumoto et al.,. 2006), and patients cannot use sentence context to compensate for the impact of abnormal tones on sentence meaning (Wang et al., 2018). However, existing research still has limitations. For instance, participant selection continues to suffer from the issues mentioned in segmental studies. Only Todd et al. (2008) controlled for illness duration, thereby discovering that pitch-change-evoked MMN could serve as an index of auditory cortex pathology. Todd et al. (2008) also used audiometric assessment to exclude participants with hearing impairment, effectively controlling for hearing loss effects on speech recognition. Regarding experimental stimuli, perception of Mandarin tones has attracted some researchers' attention, but compared to research on non-tonal languages, studies on Mandarin tones remain in their infancy and warrant greater attention from Chinese scholars. In terms of experimental design, Wang et al. (2017) conducted both behavioral and neuroscientific experiments on the same group of participants and combined the results to identify P300 as the neural basis of impaired speech processing in schizophrenia. Their research not only identified speech processing deficits but also revealed the underlying neural mechanisms, fully utilizing the advantages of combining both experimental approaches.

3.2 Duration and Intensity Perception

Duration refers to the length of time a sound lasts from beginning to end. As an attribute of sound, duration cannot be separated from the broader temporal context of speech production (Clark & Yallop, 1990). Therefore, research on duration perception in schizophrenia patients is typically linked to temporal processing. Thoenes and Oberfeld (2017) employed meta-analysis to examine visual and auditory temporal perception abilities in schizophrenia patients, revealing lower accuracy in time perception.

Research on duration perception can be divided into two categories based on experimental stimuli. The first involves different intervals between sound stimuli. For example, Davalos, Kisley, and Ross (2003) found that schizophrenia patients had difficulty discriminating intervals between non-speech sounds, indicating abnormal auditory temporal perception. Neuroscientific experiments such as Shelley et al. (1991) found that, compared to healthy controls, schizophrenia patients showed smaller MMN responses to both lengthened and shortened stimulus intervals, with the reduction being more pronounced for lengthened intervals. Erwin, Mawhinney-Hee, Gur, and Gur (1991) found that stimuli with different intervals produced smaller amplitude variation ranges in patients' P1 and N1 responses. Since P1 is generated in the thalamus, this result is consistent with reports of thalamic abnormalities in patients.

The second category involves different lengths of sound stimuli. For instance, Davalos, Rojas, and Tregellas (2011) found that schizophrenia patients showed poorer performance than healthy controls in identifying the length of non-speech sounds. Moreover, fMRI recordings revealed weaker activation in the striatum and insula in schizophrenia patients compared to healthy controls when identifying sound length, suggesting possible damage in these regions. Atkinson, Michie, and Schall (2012) found that MMN elicited by duration changes was significantly reduced in short-illness-duration schizophrenia patients compared to healthy controls, and that this condition existed before illness onset. Shin et al. (2009) (using MEG) and Atkinson et al. (2012) (using EEG) investigated whether duration-change-evoked MMN could serve as a predictive marker for schizophrenia, both finding that individuals at ultra-high risk for schizophrenia showed significantly smaller MMN amplitudes or MMNm dipole moments in response to duration deviants compared to healthy controls. Furthermore, reduced MMN amplitude was significantly correlated with decreased gray matter volume in the left Heschl' s gyrus (Salisbury, Kuroki, Kasai, Shenton, & McCarley, 2007), suggesting that gray matter reduction in certain brain regions may also contribute to speech perception deficits. Atkinson et al. (2012) proposed that duration-deviant-evoked MMN could serve as a marker of prodromal schizophrenia symptoms, meaning MMN could be used to measure the likelihood of ultra-high-risk individuals eventually developing schizophrenia. In addition to MMN, P3a is also an indicator for predicting onset in ultra-high-risk populations. Atkinson et al. (2012) found that P3a amplitude was reduced but not significantly in ultra-high-risk individuals, while first-episode psychosis patients showed significantly smaller P3a amplitudes. Moreover, the trends for MMN and P3a were not correlated, suggesting these two indicators reflect different types of impairment.

Intensity typically refers to the magnitude of sound pressure variations in speech signals (Clark & Yallop, 1990). Todd et al. (2008) examined performance in discriminating pitch, duration, and intensity of non-speech sounds in healthy controls and schizophrenia patients with average illness durations of 2.6 years and 18.9 years. They found that both long- and short-illness-duration patients showed larger intensity discrimination thresholds than healthy controls, meaning patient groups required greater intensity differences to distinguish sounds, though this difference was significant only for short-illness-duration patients compared to controls. Additionally, when examining MMN evoked by intensity changes, they found that short-illness-duration patients produced significantly smaller MMN amplitudes than healthy controls, while long-illness-duration patients showed no difference from healthy individuals. The authors analyzed that since intensity-change-evoked MMN amplitudes decrease with age in healthy controls, intensity-change-evoked MMN may exhibit reduced sensitivity in long-illness-duration patients.

Additionally, Gudlowski et al. (2009) and Park, Lee, Kim, and Bae (2010) found that sounds of different intensities produced smaller variation ranges in the loudness-dependence of auditory evoked potential (LDAEP) in schizophrenia patients. LDAEP represents amplitude changes in N1 and P2 responses elicited by variations in sound stimulus intensity, including loudness. In normal individuals, LDAEP amplitude increases with sound intensity or loudness. Teichert (2017) considered blunted LDAEP a marker of altered auditory function in schizophrenia patients, believed to result from altered serotonin function, as serotonin is a neurotransmitter produced in the body.

Beyond auditory speech processing deficits, numerous studies have examined speech processing in schizophrenia from the perspective of dyslexia mechanisms, finding that reading difficulties in schizophrenia are associated with impaired speech processing abilities (Whitford, O' Driscoll, & Titone, 2018). For example, Revheim et al. (2014) found that poor reading ability in schizophrenia patients was highly correlated with poor visual and auditory perception abilities, with auditory perception measured using MMN evoked by pitch, duration, and intensity changes. However, clinical high-risk individuals for schizophrenia showed only visual perception deficits, leaving uncertain whether speech processing impairments in reading exist before illness onset. Carrión et al. (2015) investigated clinical high-risk individuals and found that, compared to healthy individuals, the clinical high-risk group showed significantly smaller MMN amplitudes in response to pitch, duration, and intensity changes in non-speech sounds. Moreover, smaller MMN amplitudes were associated with poor reading ability, slow processing speed, and poor social and role functioning in the clinical high-risk group, suggesting that primary auditory processing deficits may lead to impairments in higher-level social abilities.

These studies demonstrate that schizophrenia patients exhibit impairments in perceiving both duration and intensity. Moreover, neuroscientific experiments have found that these impairments may be associated with damage to specific brain regions and components, such as the thalamus, striatum, insula, left Heschl' s gyrus, and the neurotransmitter serotonin. Additionally, by examining the relationship between reading disabilities and speech processing, primary auditory processing deficits may lead to impairments in higher-level social abilities such as reading. Furthermore, duration-deviant-evoked MMN and P3a can be used to measure the likelihood of ultra-high-risk individuals eventually developing schizophrenia, providing valuable reference for clinical diagnosis. However, these studies also have limitations. Regarding participant selection, the aforementioned issues persist, though existing research (e.g., Atkinson et al., 2012) has expanded the participant pool to include ultra-high-risk populations, exploring pre-illness symptom manifestations that could aid diagnostic assistance. Regarding experimental stimuli, existing studies have examined duration and intensity perception using only non-speech sounds, lacking specific investigation of duration and intensity in common speech sounds. This is particularly relevant for Mandarin Chinese as a tonal language, where duration plays a crucial role in tone recognition (Liu & Samuel, 2004) and directly affects speech perception and comprehension.

3.3 Emotional Prosody Perception

Emotion perception is an aspect of social cognition. Once emotion perception is impaired, misunderstandings and inappropriate social responses may occur (Hoekert et al., 2007). Therefore, correctly perceiving emotions is crucial for understanding speakers' emotional states and intentions. There are two pathways for emotion perception: observing facial expressions and processing emotional prosody information conveyed through speech, with the latter being the focus of this section.

Currently, most research on emotional prosody perception in schizophrenia patients has focused on native English speakers. Chan, Wong, Wang, and Lee (2008) investigated emotional prosody perception in Chinese schizophrenia patients and examined neuropsychological predictors of emotion recognition ability. After assessing remitted paranoid and non-paranoid schizophrenia patients and healthy individuals on emotional prosody recognition tasks and neuropsychological tests measuring attention and visual perception, they found that remitted non-paranoid patients exhibited emotional prosody perception deficits, while remitted paranoid patients performed similarly to healthy controls. This indicates that emotional recognition impairments persist in remitted schizophrenia patients. Additionally, neuropsychological tests revealed that paranoid schizophrenia patients had deficits in interference control, while non-paranoid schizophrenia patients showed attention mechanism impairments.

In research on emotional prosody perception in patients, both ERP and MRI results have demonstrated auditory cortex damage. Pinheiro et al. (2014) recorded event-related potentials in schizophrenia patients and healthy controls to investigate (1) the temporal course of emotional prosody processing and (2) the influence of prosodic and semantic cues on emotional prosody processing. The stimuli were prosodic words presented in two conditions: with both prosody and semantics (SCC) and with prosody but without semantics (PPC). The results showed that, compared to healthy controls, the patient group exhibited reduced P50 when processing happy PPC words, reduced N100 when processing neutral SCC words, neutral PPC words, and emotional SCC words, and increased P200 when processing happy SCC words. Behavioral results showed higher error rates when patients processed angry SCC words and happy PPC words. Since P50 is an index of sensory memory formation in the primary auditory cortex, N100 is associated with early auditory encoding and primarily generated in the superior temporal gyrus, and P200 is mainly generated in temporal cortex regions such as the planum temporale, these abnormal indicators all reflect auditory cortex damage in patients. The authors concluded that both sensory processing abnormalities and higher-level processing deficits contribute to emotional prosody processing impairments in schizophrenia, with the severity depending on stimulus complexity.

Leitman et al. (2007) used diffusion tensor imaging (DTI) to analyze the neural mechanisms underlying prosodic processing abnormalities in schizophrenia. After examining patient and healthy control groups on emotional and non-emotional prosody perception tasks, they found that patients exhibited impairments in perceiving both emotion and semantics in language, such as in distinguishing statements from questions and happy from sad speech. DTI results showed structural and functional damage in patients' primary auditory cortex. Specifically, patients showed lower fractional anisotropy in brain regions radiating primary auditory abilities (from the medial geniculate body of the thalamus to Heschl' s gyrus and dorsal and ventral auditory pathways), with lower fractional anisotropy thought to reflect damage to axons or myelin sheaths (which are essential for signal transmission within axons) (Kubicki et al., 2007; Owen et al., 2016). Myelin damage causing speech perception problems has been confirmed by many studies. For instance, abnormal central nervous system myelin can lead to auditory processing disorders, and different myelination patterns in auditory neuron axons can affect conduction velocities of high- and low-frequency sound signals, thereby impacting sound localization ability —which is crucial for speech perception in noisy, complex environments. Additionally, abnormal central nervous system myelin may be an important cause of sound hypersensitivity and auditory perceptual distortion in autism spectrum disorder patients (Ford et al., 2015; Long, Wan, Robert, & Corfas, 2018).

Mitchell, Elliott, Barry, Cruttenden, and Woodruff (2004) used fMRI to monitor participants' performance during passive listening and active emotional prosody perception. They found that schizophrenia patients showed normal right-hemisphere lateralization when passively perceiving pure emotional prosody but displayed left-hemisphere lateralization when perceiving unfiltered emotional prosody. When patients actively perceived emotional prosody, they showed greater activation in the left insula, suggesting left-hemisphere lateralization in emotional prosody perception that differs from the normal right-hemisphere lateralization. The authors analyzed that this might result from reduced insular volume in patients. Additionally, the results showed functional impairments in patients' superior and middle temporal gyri.

Wylie and Tregellas (2010) summarized research on emotional prosody in schizophrenia, finding that: (1) in terms of onset timing, emotional prosody impairments are already present in first-episode schizophrenia patients; (2) in terms of scope, schizophrenia patients exhibit impairments in both perceiving and expressing emotional prosody. Specifically, schizophrenia patients with poor emotional prosody perception performance show response times three times longer than normal individuals (Hoekert et al., 2007). (3) In terms of level, patients show impairments in perceiving emotional prosody in sentences, words, syllables, and non-syllabic sounds. (4) In terms of severity, prosodic impairments are more pronounced when discriminating negative emotions such as anger, fear, and sadness. (5) Regarding relationships with other deficits, patients' emotional prosody discrimination ability correlates with negative symptom severity, semantic processing ability, and basic pitch perception ability. Leitman et al. (2008) found that impaired emotional prosody perception in schizophrenia was related to impaired basic pitch perception ability, with results showing that patients' ability to utilize pitch-based acoustic signals was significantly poorer, suggesting that inability to process key acoustic features in prosody may cause emotional prosody perception deficits. Kantrowitz et al. (2013) further investigated the relationship between emotional prosody perception and basic pitch perception. In their experiment, schizophrenia patient and healthy control groups listened to 38 synthesized frequency-modulated tones that simulated major acoustic features of human speech. Participants were asked to identify five emotions in these tones, while researchers analyzed the contributions of mean fundamental frequency and its variability and the presence of high-frequency energy to emotion identification. Additionally, participants completed pitch-matching experiments and vocal emotion recognition tasks. They found that identification patterns for frequency-modulated tones correlated with basic pitch identification ability, vocal emotion recognition ability, and negative symptoms.

In summary, research on emotional prosody perception in schizophrenia is relatively extensive, providing in-depth understanding of various aspects of emotional prosody perception deficits and revealing the relationship between emotional prosody perception and basic pitch perception, demonstrating that primary auditory cognitive abilities influence higher-level emotional cognition. Moreover, neuroscientific research examining the three indicators of P50, N100, and P200 has also identified auditory cortex abnormalities when patients perceive emotional prosody, corroborating conclusions from segmental and linguistic prosody studies that found auditory cortex abnormalities in patients. In fact, beyond expressing semantics and emotion, voices can provide much implicit information, such as speaker identity cues. Chhabra, Badcock, Maybery, and Leung (2012) found that schizophrenia patients could distinguish voices based on mean fundamental frequency just like normal individuals, indicating that patients' ability to use pitch-based cues for speaker identification remains intact. However, both schizophrenia patients with and without auditory hallucinations rarely use formant dispersion to identify voices, whereas healthy controls use formant dispersion to perceive speakers' dominance tendencies, gender, weight, and age. This suggests that schizophrenia patients exhibit abnormalities in processing identity information in voices.

4. Summary and Future Directions

In summary, current research has examined speech perception performance across segmental and suprasegmental levels and various aspects in schizophrenia patients. Behavioral experiments have found that, compared to healthy individuals, patients exhibit impairments in perceiving consonants and vowels at the segmental level and in linguistic and emotional prosody at the suprasegmental level. Neuroscientific experiments have further explained that speech perception deficits may result from damage to specific brain regions and components. Specifically, numerous studies examining indicators such as MMN, P50, N100, and P200 have found auditory cortex damage in patients. Additionally, research (e.g., Davalos et al.„ 2011; Erwin et al., 1991; Teichert, 2017) has identified damage in other brain regions or components, including the striatum, insula, thalamus, and the neurotransmitter serotonin. Placing these brain abnormalities within the auditory system reveals how they affect patients' speech perception. Human speech perception relies on collaborative functioning across auditory system components. The neural portion of the mammalian auditory system roughly includes inner and outer hair cells of the cochlea, spiral ganglion neurons, the cochlear branch of the vestibulocochlear nerve (auditory nerve), cochlear nucleus complex, trapezoid body, superior olivary complex, lateral lemniscus, inferior colliculus, brachium of inferior colliculus, medial geniculate body, thalamocortical auditory radiations, and cerebral auditory cortex (Webster, Popper, & Fay, 1992). According to current research on schizophrenia patients, damage to any of these regions can cause auditory abnormalities. In mammals, all known auditory information enters the brain through the auditory nerve (Webster et al., 1992). Serotonin abnormalities in patients demonstrate disruptions in this transmission, while thalamic and auditory cortex abnormalities reflect impairments in higher-level auditory processing within the auditory system. Additionally, research (e.g., Davalos et al.„ 2011) has found damage to patients' striatum and insula, which is intimately connected to these regions' locations and functions in the brain. For instance, the striatum connects to the thalamus and receives auditory information transmitted from the thalamus, while the insula connects to the auditory cortex and is closely related to language function (Price, 2000; Webster et al., 1992).

Overall, current research on speech perception in schizophrenia patients has achieved certain accomplishments. However, beyond the limitations in exploring various perceptual levels mentioned above, several aspects warrant attention in future research.

First, although existing research has identified numerous brain region damages that can explain speech perception abnormalities, this understanding remains insufficiently deep, and the functional relationships between brain regions are still unclear, requiring further expansion and investigation. Moreover, speech perception is a fundamental human cognitive function. Primary speech perception abnormalities, such as pitch perception deficits, may cause errors in lexical and sentence perception and impair higher-level cognitive abilities like emotion perception and reading, thereby affecting language communication. Therefore, the relationship between speech perception abilities and other cognitive functions in patients requires greater attention, and researchers should emphasize using advanced neuroscientific methods to deeply investigate these two aspects. As technology develops and research conditions improve, techniques applied to speech perception exploration in schizophrenia patients will become more diversified. Beyond commonly used methods such as ERP, EEG, and MRI, technologies like MEG (magnetoencephalography) and fNIRS (functional near-infrared spectroscopy) are increasingly being applied in speech research. For example, Kasai et al. (2003) used MEG to investigate phoneme perception issues in schizophrenia patients, 杨海波, 张雪健, 周菘, 刘颖, and 白学军 (2014) summarized research progress on fNIRS in speech processing, Beversdorf, Metzger, Nelson, Alonso, and Kight (1995) used SPECT to examine word perception in normal individuals, and Talavage, Gonzalez-Castillo, and Scott (2014) summarized PET applications in auditory speech perception. Widespread application of these technologies will undoubtedly deepen understanding of the mysteries of speech processing in schizophrenia patients.

Second, speech systems differ across languages, reflected at both segmental and suprasegmental levels and aspects. Listeners with different native language backgrounds also differ in their attentional focus and processing strategies for relevant acoustic information (Strange & Shafer, 2008), with varying processing difficulty. Therefore, systematic comparison of speech processing performance in patients with different native language backgrounds is necessary to better identify commonalities and individualities in cognitive processing of speech perception in mental disorders. Existing research on speech perception in schizophrenia has mostly used native English speakers, with fewer studies on other languages worldwide. Since each language has a different speech system, research on schizophrenia patients with different native language backgrounds is needed, with particular emphasis on increasing speech research on Chinese schizophrenia patients. The reasons are as follows: First, 20% of schizophrenia patients worldwide speak tonal languages like Mandarin, yet most research on speech perception in schizophrenia has focused on non-tonal language speakers (Yang et al., 2012). This situation not only mismatches the important role of tone in speech perception but also neglects the large population of tonal language-speaking schizophrenia patients. Although some studies (e.g., Tsao et al., 2013; Wang et al., 2017; Wang et al., 2018; Yang et al., 2012) have filled gaps in research on tone processing in schizophrenia patients, current research remains in its early stages. Since Mandarin is the most widely spoken tonal language in the world, studying tonal languages through Chinese schizophrenia patients is significant. Second, alphabetic scripts like English follow grapheme-to-phoneme conversion rules, using letters as visual symbols to map spoken phonemes. Therefore, letter words are primarily read through assembled phonology—combining phonemes. In contrast, Chinese uses characters as basic writing units, with character strokes having no phoneme-corresponding components. Thus, understanding Chinese characters' phonological encoding requires addressed phonology—searching for phonological information stored in the cognitive system. This difference in orthography-phonology mapping between the two languages may lead to different neural mechanisms (Tan, Laird, Li, & Fox, 2005). Additionally, Mandarin has four tones and many homophonic characters with different meanings, placing higher cognitive demands on accurate processing of both phonemes and tones (Yang et al, 2015). These characteristics differ substantially from alphabetic scripts. Therefore, language features of Chinese schizophrenia patients may not completely align with research findings from Western countries. To more clearly understand language characteristics and their neural mechanisms in Chinese schizophrenia patients and provide effective assistance for diagnosis and etiological mechanism discovery in Chinese patients, extensive Mandarin research is essential. Finally, China has unique cultural and social customs, and how corresponding social-emotional factors interact with speech functions to affect speech performance in mental disorder patients requires more research.

Third, most existing research examining speech perception in schizophrenia patients has been conducted in quiet laboratory conditions, disconnected from the noisy living environments of real life. Therefore, researchers need to pay attention to speech perception in schizophrenia patients under noisy conditions. When perceiving speech in noise, listeners must separate the speaker's speech signal from various mixed sound waves to identify target sounds, producing the "cocktail party effect." Many studies have found that in noisy backgrounds, speech perception in certain special populations is more susceptible to noise masking effects than in normal populations, such as older adults (Nagaraj, 2017) and second language learners (徐灿 et al., 2018; 杨小虎 & 赵勇, 2014). Do schizophrenia patients also experience greater difficulties? Do patients with different symptom profiles show different degrees of impact from different types of noise and signal-to-noise ratios? Additionally, many schizophrenia patients experience auditory hallucinations, and several studies (e.g., Tsao et al., 2013; Wang et al., 2018) have found that schizophrenia patients with hallucination symptoms perform worse on speech perception tasks than those without hallucinations. However, the mechanisms and reasons through which hallucinations, as a form of internal noise, affect patients' speech perception have not been thoroughly investigated. Therefore, examining these issues has important theoretical and practical significance and warrants attention in future research.

References

曹剑芬. (2002). 汉语声调与语调的关系. 中国语文 (3), 195-202.
华维芬. (1993). 英语超音段音位浅析. 外语研究 (4), 44-48.
徐灿, 杨小虎, 汪玉霞, 张辉, 丁红卫, 刘畅. (2018). 语音型噪音对二语者汉语元音声调感知的影响. 心理与行为研究, 16(1), 22-30.
杨海波, 张雪健, 周菘, 刘颖, 白学军. (2014). 语音加工的功能性近红外脑成像研究进展. 心理与行为研究, 12(4), 566-571.
杨小虎, 赵勇. (2014). 噪音背景对二语语音感知的影响. 心理科学进展, 22(6), 934-942.

Atkinson, R. J., Michie, P. T., & Schall, U. (2012). Duration mismatch negativity and P3a in first-episode psychosis and individuals at ultra-high risk of psychosis. *Biological Psychiatry, 71*(2), 98-104.

Bellani, M., Perlini, C., & Brambilla, P. (2009). Language disturbances in schizophrenia. *Epidemiology and Psychiatric Sciences, 18*(4), 314-317.

Beversdorf, D., Metzger, S., Nelson, D., Alonso, R., & Kight, J. (1995). Single-word auditory stimulation and regional cerebral blood flow as studied by SPECT. *Psychiatry Research: Neuroimaging, 61*(3), 181-189.

Carrión, R. E., Cornblatt, B. A., McLaughlin, D., Chang, J., Auther, A. M., Olsen, R. H., & Javitt, D. C. (2015). Contributions of early cortical processing and reading ability to functional status in individuals at clinical high risk for psychosis. *Schizophrenia Research, 164*(1-3), 1-7.

Chan, C. C., Wong, R., Wang, K., & Lee, T. M. (2008). Emotion recognition in Chinese people with schizophrenia. *Psychiatry Research, 157*(1), 67-76.

Chhabra, S., Badcock, J. C., Maybery, M. T., & Leung, D. (2012). Voice identity discrimination in schizophrenia. *Neuropsychologia, 50*(12), 2730-2735.

Cienfuegos, A., March, L., Shelley, A. M., & Javitt, D. C. (1999). Impaired categorical perception of synthetic speech sounds in schizophrenia. *Biological Psychiatry, 45*(1), 82-88.

Clark, J., & Yallop, C. (1990). *An introduction to phonetics and phonology.* Oxford, England: Blackwell.

Covington, M. A., He, C., Brown, C., Naçi, L., McClain, J. T., Fjordbak, B. S., ···Brown, J. (2005). Schizophrenia and the structure of language: The linguist's view. *Schizophrenia Research, 77*(1), 85-98.

Davalos, D. B., Kisley, M. A., & Ross, R. G. (2003). Effects of interval duration on temporal processing in schizophrenia. *Brain and Cognition, 52*(3), 295-301.

Davalos, D. B., Rojas, D. C., & Tregellas, J. R. (2011). Temporal processing in schizophrenia: Effects of task-difficulty on behavioral discrimination and neuronal responses. *Schizophrenia Research, 127*(1-3), 123-130.

Dondé, C., Luck, D., Grot, S., Leitman, D. I., Brunelin, J., & Haesebaert, F. (2017). Tone-matching ability in patients with schizophrenia: A systematic review and meta-analysis. *Schizophrenia Research, 181*, 94–99.

Edwards, J., Jackson, H. J., & Pattison, P. E. (2002). Emotion recognition via facial expression and affective prosody in schizophrenia: A methodological review. *Clinical Psychology Review, 22*(6), 789–832.

Erwin, R. J., Mawhinney-Hee, M., Gur, R. C., & Gur, R. E. (1991). Midlatency auditory evoked responses in schizophrenia. *Biological Psychiatry, 30*(5), 430–442.

Ford, M. C., Alexandrova, O., Cossell, L., Stange-Marten, A., Sinclair, J., Kopp-Scheinpflug, C., ···Grothe, B. (2015). Tuning of Ranvier node and internode properties in myelinated axons to adjust action potential timing. *Nature Communications, 6*, 8073.

Gudlowski, Y., Özgürdal, S., Witthaus, H., Gallinat, J., Hauser, M., Winter, C., ···Juckel, G. (2009). Serotonergic dysfunction in the prodromal, first-episode and chronic course of schizophrenia as assessed by the loudness dependence of auditory evoked activity. *Schizophrenia Research, 109*(1–3), 141–147.

Hoekert, M., Kahn, R. S., Pijnenborg, M., & Aleman, A. (2007). Impaired recognition and expression of emotional prosody in schizophrenia: Review and meta-analysis. *Schizophrenia Research, 96*(1–3), 135–145.

Hugdahl, K., Løberg, E. M., Falkenberg, L. E., Johnsen, E., Kompus, K., Kroken, R. A., ···Özgören, M. (2012). Auditory verbal hallucinations in schizophrenia as aberrant lateralized speech perception: Evidence from dichotic listening. *Schizophrenia Research, 140*(1–3), 59–64.

Javitt, D. C. (2000). Intracortical mechanisms of mismatch negativity dysfunction in schizophrenia. *Audiology and Neurotology, 5*(3–4), 207–215.

Javitt, D. C., Spencer, K. M., Thaker, G. K., Winterer, G., & Hajós, M. (2008). Neurophysiological biomarkers for drug development in schizophrenia. *Nature Reviews Drug Discovery, 7*(1), 68–83.

Kantrowitz, J. T., Leitman, D. I., Lehrfeld, J. M., Laukka, P., Juslin, P. N., Butler, P. D., ···Javitt, D. C. (2013). Reduction in tonal discriminations predicts receptive emotion processing deficits in schizophrenia and schizoaffective disorder. *Schizophrenia Bulletin, 39*(1), 86–93.

Kasai, K., Nakagome, K., Itoh, K., Koshida, I., Hata, A., Iwanami, A., ···Kato, N. (2002). Impaired cortical network for preattentive detection of change in speech sounds in schizophrenia: A high-resolution event-related potential study. *American Journal of Psychiatry, 159*(4), 546–553.

Kasai, K., Yamada, H., Kamio, S., Nakagome, K., Iwanami, A., Fukuda, M., ···Kato, N. (2003). Neuromagnetic correlates of impaired automatic categorical perception of speech sounds in schizophrenia. *Schizophrenia Research, 59*(2–3), 159–172.

Kubicki, M., McCarley, R., Westin, C. F., Park, H. J., Maier, S., Kikinis, R., ···Shenton, M. E. (2007). A review of diffusion tensor imaging studies in schizophrenia. *Journal of Psychiatric Research, 41*(1-2), 15-30.

Kugler, B. T., & Caudrey, D. J. (1983). Phoneme discrimination in schizophrenia. *British Journal of Psychiatry, 142*(1), 53-59.

Ladefoged, P., & Johnson, K. (2011). *A course in phonetics.* Boston, MA: Wadsworth.

Leitman, D. I., Hoptman, M. J., Foxe, J. J., Saccente, E., Wylie, G. R., Nierenberg, J., ···Javitt, D. C. (2007). The neural substrates of impaired prosodic detection in schizophrenia and its sensorial antecedents. *American Journal of Psychiatry, 164*(3), 474-482.

Leitman, D. I., Laukka, P., Juslin, P. N., Saccente, E., Butler, P., & Javitt, D. C. (2008). Getting the cue: Sensory contributions to auditory emotion recognition impairments in schizophrenia. *Schizophrenia Bulletin, 36*(3), 545-556.

Liu, S., & Samuel, A. G. (2004). Perception of Mandarin lexical tones when F0 information is neutralized. *Language and Speech, 47*(2), 109-138.

Long, P., Wan, G., Roberts, M. T., & Corfas, G. (2018). Myelin development, plasticity, and pathology in the auditory system. *Developmental Neurobiology, 78*(2), 80-92.

Matsumoto, K., Samson, G. T., O' daly, O. D., Tracy, D. K., Patel, A. D., & Shergill, S. S. (2006). Prosodic discrimination in patients with schizophrenia. *The British Journal of Psychiatry, 189*(2), 180-181.

Mitchell, R. L., Elliott, R., Barry, M., Cruttenden, A., & Woodruff, P. W. (2004). Neural response to emotional prosody in schizophrenia and in bipolar affective disorder. *The British Journal of Psychiatry, 184*(3), 223-230.

Nagaraj, N. K. (2017). Working memory and speech comprehension in older adults with hearing impairment. *Journal of Speech, Language, and Hearing Research, 60*(10), 2949-2964.

Owen, M. J., Sawa, A., & Mortensen, P. B. (2016). Schizophrenia. *The Lancet, 388*(10039), 86-97.

Park, Y. M., Lee, S. H., Kim, S., & Bae, S. M. (2010). The loudness dependence of the auditory evoked potential (LDAEP) in schizophrenia, bipolar disorder, major depressive disorder, anxiety disorder, and healthy controls. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 34*(2), 313-316.

Pinheiro, A. P., Rezaii, N., Rauber, A., Liu, T., Nestor, P. G., McCarley, R. W., ···Niznikiewicz, M. A. (2014). Abnormalities in the processing of emotional prosody from single words in schizophrenia. *Schizophrenia Research, 152*(1), 235-241.

Price, C. (2000). The anatomy of language: Contributions from functional neuroimaging. *Journal of Anatomy, 197*(3), 335-359.

Revheim, N., Corcoran, C. M., Dias, E., Hellmann, E., Martinez, A., Butler, P. D., ···Javitt, D. C. (2014). Reading deficits in schizophrenia and individuals at high clinical risk: Relationship to sensory function, course of illness, and psychosocial outcome. *American Journal of Psychiatry, 171*(9), 949-959.

Salisbury, D. F., Kuroki, N., Kasai, K., Shenton, M. E., & McCarley, R. W. (2007). Progressive and interrelated functional and structural evidence of post-onset brain reduction in schizophrenia. *Archives of General Psychiatry, 64*(5), 521-529.

Shelley, A. M., Ward, P. B., Catts, S. V., Michie, P. T., Andrews, S., & McConaghy, N. (1991). Mismatch negativity: An index of a preattentive processing deficit in schizophrenia. *Biological Psychiatry, 30*(10), 1059-1062.

Shin, K. S., Kim, J. S., Kang, D. H., Koh, Y., Choi, J. S., O' Donnell, B. F., ···Kwon, J. S. (2009). Pre-attentive auditory processing in ultra-high-risk for schizophrenia: A magnetoencephalography study. *Biological Psychiatry, 65*(12), 1071-1078.

Strange, W., & Shafer, V. L. (2008). Speech perception in second language learners: The re-education of selective perception. In J. G. H. Edwards & M. L. Zampini (Vol. Eds.), *Studies in bilingualism: Vol. 36. Phonology and second language acquisition* (pp. 153-191). Amsterdam: John Benjamins.

Sweet, R. A., Bergen, S. E., Sun, Z., Marcsisin, M. J., Sampson, A. R., & Lewis, D. A. (2007). Anatomical evidence of impaired feedforward auditory processing in schizophrenia. *Biological Psychiatry, 61*(7), 854-864.

Talavage, T. M., Gonzalez-Castillo, J., & Scott, S. K. (2014). Auditory neuroimaging with fMRI and PET. *Hearing Research, 307*, 4-15.

Tan, L. H., Laird, A. R., Li, K., & Fox, P. T. (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: A meta-analysis. *Human Brain Mapping, 25*(1), 83-91.

Teichert, T. (2017). Loudness- and time-dependence of auditory evoked potentials is blunted by the NMDA channel blocker MK-801. *Psychiatry Research, 256*, 202-206.

Thoenes, S., & Oberfeld, D. (2017). Meta-analysis of time perception and temporal processing in schizophrenia: Differential effects on precision and accuracy. *Clinical Psychology Review, 54*, 44-64.

Todd, J., Michie, P. T., Schall, U., Karayanidis, F., Yabe, H., & Näätänen, R. (2008). Deviant matters: Duration, frequency, and intensity deviants reveal different patterns of mismatch negativity reduction in early and late schizophrenia. *Biological Psychiatry, 63*(1), 58-64.

Tsao, F. M., Chiang, S. K., & Liu, H. M. (2013). Lexical tone and consonant perception in subtypes of schizophrenia. *Journal of the Acoustical Society of America, 134*(5), 4235.

van Os, J., Kenis, G., & Rutten, B. P. (2010). The environment and schizophrenia. *Nature, 468*(7321), 203–212.

Wang, J., Liu, Q., Wydell, T. N., Liao, J., Wang, F., Quan, W., ···Dong, W. (2017). Electrophysiological basis of reading related phonological impairment in Chinese speakers with schizophrenia: An ERP study. *Psychiatry Research: Neuroimaging, 261*, 65–71.

Wang, J., Wydell, T. N., Zhang, L., Quan, W., Tian, J., Liu, J., & Dong, W. (2018). The underlying mechanism of deficits of speech comprehension and hallucinations in Chinese patients with schizophrenia. *Journal of Psychiatric Research, 97*, 16–21.

Webster, D.B., Popper, A.N., & Fay, R.R. (Eds). (1992). *The mammalian auditory pathway: Neuroanatomy*. New York: Springer-Verlag.

Whitford, V., O' Driscoll, G. A., & Titone, D. (2018). Reading deficits in schizophrenia and their relationship to developmental dyslexia: A review. *Schizophrenia Research, 193*, 11–22.

Wylie, K. P., & Tregellas, J. R. (2010). The role of the insula in schizophrenia. *Schizophrenia Research, 123*(2–3), 93–104.

Yang, L., Chen, S., Chen, C. M., Khan, F., Forchelli, G., & Javitt, D. C. (2012). Schizophrenia, culture and neuropsychology: Sensory deficits, language impairments and social functioning in Chinese-speaking schizophrenia patients. *Psychological Medicine, 42*(7), 1485–1494.

Yang, X., Wang, Y., Xu, L., Zhang, H., Xu, C., & Liu, C. (2015). Aging effect on Mandarin Chinese vowel and tone identification. *The Journal of the Acoustical Society of America, 138*(4), EL411–EL416.

Note: Figure translations are in progress. See original paper for figures.