ANNUAL REVIEWS

Annual Review of Linguistics

Speech Prosody in Mental Disorders

Hongwei Ding[1] and Yang Zhang[2]

[1]Speech-Language-Hearing Center, School of Foreign Languages, Shanghai Jiao Tong University, Shanghai, China; email: hwding@sjtu.edu.cn

[2]Department of Speech-Language-Hearing Sciences and Masonic Institute for the Developing Brain, University of Minnesota, Minneapolis, Minnesota, USA

Annu. Rev. Linguist. 2023. 9:335–55

First published as a Review in Advance on October 24, 2022

The *Annual Review of Linguistics* is online at linguistics.annualreviews.org

https://doi.org/10.1146/annurev-linguistics-030421-065139

Keywords

mental disorders, speech prosody, prosody recognition, prosody production, pragmatic prosody, affective prosody

Abstract

In response to uncovering brain mechanisms underlying vocal communication and searching for biomarkers for mental illnesses, speech prosody has been increasingly studied in recent years in multiple disciplines, including psycholinguistics. In this article, we provide an up-to-date synthesis of the theoretical foundation and empirical evidence to profile linguistic and emotional prosody in the proper characterization of mental disorders, including schizophrenia, autism, Alzheimer's disease, and depression. Our review reveals a need to develop theoretically motivated and methodologically integrated approaches to the study of context-driven comprehension and expression of pragmatic-affective prosody, which will help elucidate the core features of socio-communicative problems in individuals with mental disorders. We propose that comprehensive models within and across the conventional cognition-emotion-language trichotomy need to be developed to integrate current findings and guide future research. In particular, there needs to be due emphasis on investigating multisensory and cross-modal effects in normal and pathological prosody research. Our review calls for multidisciplinary efforts to address the challenging issues to inform and inspire the advancement of linguistic theories and psychiatric diagnosis and treatment.

1. INTRODUCTION

Speech prosody, which functions as a bridge between verbal behavior and mental disorders, has been increasingly recognized as a topic of significant interest in psychiatric research in recent decades (see Icht et al. 2021, Lin et al. 2018, Low et al. 2020, Robin et al. 2020, Zhang et al. 2021 for reviews). In this review, we discuss the theoretical foundations and current findings as well as controversial issues in the study of speech prosody in mental disorders. We highlight the need to integrate emotion and cognition with language in this interdisciplinary research field to broaden the horizons of linguistics and push for new frontiers in psychiatry. To this end, the article starts with a brief introduction to mental disorders and speech prosody, followed by two separate sections of critical literature review on the perception and production of speech prosody and conclusive remarks on current perspectives and future directions.

1.1. Mental Disorders

Mental disorders (or psychiatric disorders) are generally characterized by a combination of abnormal thoughts, perceptions, emotions, behavior, and relationships with others. Major mental disorders include depression, bipolar disorder, dementia, schizophrenia, and developmental disorders such as autism (World Health Organization 2022). Schizophrenia, a severe psychiatric disorder affecting approximately 20 million people, can be destructive for both the individuals affected and society at large. It is characterized by positive symptoms (e.g., delusions, hallucinations, and disorganized speech) and negative symptoms (e.g., diminished emotional expression) (McCutcheon et al. 2020). Autism spectrum disorder (ASD), which affects more than 52 million people worldwide, is characterized by persistent impairments in social communication and interaction and by restricted and repetitive behaviors, with onset of the condition in early childhood (Baxter et al. 2015). Reported prevalence rates of ASD in many countries have been growing steadily for years, and more awareness of diagnosis and intervention at earlier ages is necessary (Baxter et al. 2015, Schreibman et al. 2015). Depression is a common mental disorder affecting more than 264 million people worldwide. It is characterized by persistent sadness and a loss of interest or pleasure in most routine activities (World Health Organization 2022). Uncertainty and isolation during the coronavirus-19 (COVID-19) pandemic have increased the prevalence of depression and anxiety, which demands more urgent attention (Wu et al. 2021). Finally, Alzheimer's disease (AD) is strongly associated with neurodegeneration and decreased cognition, including decreased language capabilities and loss of both memory and emotional stability. AD, the primary cause of dementia, affects up to 45 million individuals worldwide (Hampel et al. 2021). With the increasing life expectancy worldwide, AD is becoming a major societal concern. Mental health problems are now so common that they are among the most burdensome of all disease classes because of their high prevalence, chronicity, early-onset age, severe impairments, and healthcare costs. Nonetheless, the burden continues to grow, with significant detrimental impacts on health and social consequences in every country of the world.

The clinical interest in speech prosody research is to incorporate speech databases into early screening for mental disorders. There are currently no reliable transdiagnostic biomarkers for detecting major mental disorders (see Carvalho et al. 2020 for a review). Therefore, the question "Can we more effectively diagnose and treat complex mental disorders?" was included in a 2021 list of 125 important questions for exploration in science (*Science* 2021). As there is no known medical cure for mental illnesses, it is suggested that the most effective approach is through early and preclinical detection, which would allow for more successful cognitive-behavior therapy. But the diagnosis of mental disorders is still entirely based on psychopathological knowledge using a one-on-one diagnostic interview, which is a slow process and involves subjective judgments. Therefore, an urgent goal is to incorporate digital phenotyping to provide an objective, data-driven, and personalized approach to diagnosing and treating mental disorders. Under this backdrop, speech has emerged as a promising digital biomarker by providing a window into brain health, as many psychiatric and neurodegenerative diseases can affect speech characteristics (Robin et al. 2020). Furthermore, it has been well documented that individuals with mental disorders show the core feature of impaired emotional and pragmatic competence of language use in social contexts (Bambini et al. 2016); such impairments can be detected in their recognition and production of prosody with speech data (Mundt et al. 2012). Thus, the investigation of prosody offers an excellent venue to explore speech and language deficits as promising biomarkers for mental disorders.

1.2. Speech Prosody

In speech communication, we employ a range of cues to interpret and express the meaning of messages. How something is said may sometimes be as important as what is actually said. However, only a few studies have focused on the para- or supra-linguistic aspects of how speech is produced, which is a modest but growing area of research (Pisanski et al. 2016). These suprasegmentals are concerned with those elements of speech that are not individual phonetic segments (vowels and consonants) but are properties of syllables and larger units of speech, which may often manifest at the segmental level (Grandjean 2021). Therefore, the more general term prosody is usually employed rather than the term suprasegmentals. Speech prosody refers to the vocal modulations that accompany speech and comprises variations in fundamental frequency (also referred to as F0 and pitch), duration, intensity, and voice quality (Belyk & Brown 2014, Cutler & Pearson 2018). The prosodic speech feature plays a vital role in everyday communication by delivering the speaker's meaning and allowing the listener to interpret the intended meaning beyond the literal meaning of the words. We usually distinguish between two kinds of prosody. The most straightforward information that prosody can express is the emotional state of the speaker (Fairbanks & Pronovost 1938), which has been referred to as emotional or affective prosody. Prosody can also provide linguistic cues to augment the verbal or communicative component, which has been referred to as intrinsic or linguistic prosody (Monrad-Krohn 1947). Although having distinct functions, emotional prosody and linguistic prosody share the same acoustic parameters, and their linguistic form is equivalent (Ladd 2008). Neuroimaging studies have demonstrated that these two functions of prosody have substantial overlapping brain activation except for several divergent networks (Belyk & Brown 2014).

Although researchers in psychology and psychiatry have long conducted numerous experiments on prosody, terminology confusion and knowledge gaps in this area need to be clarified. Three categories of prosody are commonly distinguished, especially in studies of the ASD population, on the basis of these prosody types having different functions in identifying the neurological mechanism underlying the competence of speakers or listeners (Loveall et al. 2021, McCann & Peppe 2003, Paul et al. 2005, Shriberg et al. 2001).

Grammatical prosody includes cues used to signal lexical stresses, syntactic boundaries, and sentence intonation (Crystal 1986). For example, word stresses and other prosodic cues can be used (*a*) to distinguish the part of speech (e.g., "INcrease" as a noun versus "inCREASE" as a verb), (*b*) to indicate the syntactic boundaries [e.g., "(apple and orange) or banana" versus "apple and (orange or banana)"], and (*c*) to signal whether a sentence is a statement or question (e.g., "Prosody is melody." versus "Prosody is melody?"). Grammatical uses of prosody are an intrinsic part of interpreting the deep structure of the grammatical construction.

Pragmatic prosody is employed to carry social and contextual information beyond grammatical meaning. It is often the most important means to convey the speaker's communicative purposes (Ladd 1980). For example, they include contrastive stresses (e.g., "She wanted RED flowers." versus "SHE wanted red flowers."), sarcasm (e.g., saying "It's great." with a tone of disgust), and cues for turn-taking (Loveall et al. 2021). Such aspects of prosody are not obligatory but are indispensable parts of speech communication.

Affective prosody, which includes the vocal changes used for the individual's personal speech style as well as for varying social functions, is used to express the general feeling state of a speaker (Bolinger 1989). For example, emotional feelings usually include six basic emotions: anger, disgust, fear, enjoyment (happiness), sadness, and surprise (Ekman 1992) (e.g., "I can't believe it." can be expressed in neutral or one of the six basic emotions).

In studies of psychopathology, linguistic prosody has usually been regarded as being equivalent to grammatical prosody, while prosody that conveys pragmatic language functions (e.g., attitudinal prosody) has been investigated separately from linguistic prosody (Mitchell & Ross 2013). However, prosody in spoken language comprehension and expression has traditionally been studied at lexical, syntactic, and discourse levels (Cutler et al. 1997, Pierrehumbert & Hirschberg 1990). As prosodic characteristics of words can be of direct use in the pragmatic analysis of an utterance, pragmatic prosody processed at the discourse level can be integrated into the linguistic structure. Therefore, from the linguistic perspective, grammatical and pragmatic functions are represented by linguistic prosody. In contrast, emotional prosody is usually understood to display a more universal feeling (Pell et al. 2009), typically including basic emotions. Complex emotions are considered more context- and society-dependent with diverse classifications, including pride and embarrassment (Chevallier et al. 2011, Icht et al. 2021). In the Cambridge Mindreading Battery tasks for adults with and without Asperger's syndrome (AS), a taxonomy of 20 complex emotions and mental states, such as distaste and insincerity, was selected (Golan et al. 2006). Some experts claimed that such social emotions are conveyed by attitudinal functions of prosody (Mitchell & Ross 2013, Pell 2006). In recent studies, attitudinal prosody is often grouped with emotional prosody under the superordinate term of affective prosody, although most previous affective prosody research focused on simple emotional prosody (Mitchell & Ross 2013, Pell 2006). Therefore, affective (or emotional) prosody in psychological research can represent pragmatic, pure emotional, or affective language functions. This clarification of prosody categories can also help us understand behavioral performances, neurological mechanisms, and related models and theories of prosodic research in mental disorders.

As a considerable amount of evidence has accumulated to show that speech prosody may link linguistics with other disciplines such as psychiatry and neuroscience, it is necessary to examine the developments and trends that may need reflection for further advancement in this field. The review has three purposes. We aim (*a*) to summarize the characteristics of the existing studies, with the focus on the atypical behavioral performances of the mentally disordered groups; (*b*) to address the issues in the current theories and models that are employed to explain abnormal prosodic performance; and (*c*) to discuss current limitations and future opportunities. These three aspects are discussed for recognition and production of prosody in separate sections, considering that different models have been proposed for perception and production, despite efforts to integrate them into a unified one (Kroger et al. 2009). Since expressive prosody in mental disorders has increasingly been studied with data-driven approaches in computer sciences (see Cummins et al. 2015, Low et al. 2020 for reviews), our review focuses mainly on receptive prosody, with a brief discussion on expressive prosody. Finally, future orientations are provided with a summary of the current research.

2. RECOGNITION OF PROSODY IN MENTAL DISORDERS

Individuals with mental disorders have many difficulties in social functioning, which may arise from their failure to accurately recognize emotional and linguistic prosody (Chevallier et al. 2011, Diehl & Paul 2013). The earliest studies showed that schizophrenic patients were comparable in their performance with normal controls in comprehending stressed prosody, while such patients were significantly inferior to the control group in understanding emotional prosody (Murphy & Cutting 1990). Therefore, most of the literature on prosody comprehension in patients with mental disorders has focused on emotional or affective prosody. Impairments in recognition of speech prosody were first investigated in the severe disorder of schizophrenia (see Edwards et al. 2002 for a review). Similar investigations followed for autism (see McCann & Peppe 2003 for a review), AD (see Klein-Koerkamp et al. 2012 for a meta-analysis), and bipolar and depressive disorders (see Tang et al. 2022 for a review).

2.1. Behavioral Performances

The currently available assessment tools focus primarily on emotional aprosodia, whereas linguistic prosody is rarely assessed. Among the numerous validated tools that have been frequently adopted to evaluate aprosodia are the Aprosodia Battery (Ross et al. 1997), Battery of Emotional Expression and Comprehension (Cancelliere & Kertesz 1990), Florida Affect Battery (Bowers et al. 1998), Montreal Protocol for the Evaluation of Communication (Joanette et al. 2004), Assessment Battery for Communication (Angeleri et al. 2008), and Comprehensive Affect Testing System (Froming et al. 2022). Most of these validated tools are written in English, with only a few instruments developed in European languages (see Benedetti et al. 2021 for a review).

2.1.1. Recognition of emotional prosody. Many studies adopted the abovementioned standardized tools for recognition experiments, but others had to develop new auditory stimuli. The primary tasks in prosody recognition were identification and discrimination, in which the participants were required to make judgments on the basis of the prosodic information conveyed by the stimuli (see Edwards et al. 2002 for a review). Verbal stimuli could be presented either without semantic contents, such as vocal sounds, frequency-modulated tones, asyllabic sounds, and meaningless utterances, or with semantic contents in the form of syllables, words, phrases, and sentences. Some tasks employed Stroop and priming paradigms in prosody recognition to determine the influence of cognitive abilities of facilitatory and inhibitory processes (see Lin et al. 2018 for a review). The main issue of these behavior studies was to examine whether the mentally disordered groups showed worse performance than the control groups in recognition of the tested prosodic functions. Other topics included finding the relationship between recognition capabilities and related factors. Recognition accuracy was usually reported as the primary indicator of participants' performance. If necessary, response time was also recorded as a criterion. Influential factors generally included participant-related factors and experiment-related factors. If neurological studies were involved, neural mechanisms influencing impaired performance were discussed (see Zhang et al. 2021 for a review).

The current findings of the behavioral prosodic performances are different from one disorder to another but show a general tendency. Despite mixed results across studies, most of the available evidence seems to support the well-accepted finding that individuals with mental illnesses are impaired in comprehending emotional prosody, but not linguistic prosody. However, a detailed examination of the causes behind the results is necessary and can help us understand the potential issues in these studies.

For severe schizophrenic disorders, extensive research on emotional prosody recognition can be found. However, by reviewing seven studies between 1964 and 1994, Edwards et al. (2002)

suggested that only thin evidence supported emotion recognition deficits in schizophrenia due to many methodological shortcomings regarding subjects' characteristics and task variables. Many articles have subsequently been published on this topic with experimental improvements, and more evidence has been accumulated. Two meta-analyses on 17 studies between 1980 and 2007 (Hoekert et al. 2007) and on 29 studies between 2001 and 2018 (Lin et al. 2018) reported impaired processing of emotional prosody in schizophrenia with very large effect sizes. However, both cognition and semantic information were shown to be influential factors of schizophrenic patients' recognition performance. As the task difficulty increased, the linguistic (semantics) and cognitive (attention) demands also increased. Therefore, without consideration of language and cognition, impairments in emotional perception might be overestimated.

Most prosody research in ASD has focused on individuals with high-functioning autism and AS, who do not have comorbid intellectual disability (ID) but exhibit reduced social cognition in comparison to other cognitive abilities (American Psychiatric Association 1994). Due to insufficient control of stimuli, contradictory findings on receptive emotional prosody were reported by two reviewed studies published between 1980 and 2002 in ASD groups (see McCann & Peppe 2003 for review). Later, a study on 33 articles published between 1990 and 2013 reported that comprehension of emotional prosody was impaired in the ASD group (see Lartseva et al. 2015 for a review). A meta-analysis also showed impaired affective prosody recognition in ASD participants on 23 papers published between 2000 and 2019 with a moderate to large pooled effect, where only basic emotions were concerned; however, the difference was no longer significant after correction for publication bias (Zhang et al. 2021). By teasing simple and complex emotions apart, Icht et al. (2021), in a meta-analysis on 23 journal articles published between 2002 and 2020, demonstrated that ASD individuals without ID showed no significant difference from controls in the identification of simple prosodic emotions (e.g., sadness, happiness) but that these individuals performed significantly worse than controls in identifying complex prosodic emotions (e.g., envy, boredom, and calmness). There was also a high likelihood of publication bias for simple emotions, which was not observed for studies on complex emotions. The results convincingly demonstrated that ASD individuals without ID could recruit their cognitive abilities to identify the simple universal emotions but had difficulty recognizing social norms of the context- and culture-dependent complex emotions. Discriminating between these more subtle emotions demands intricate social, emotional, and linguistic faculties (Icht et al. 2021).

Studies of emotional recognition in AD have not been as extensive as those in schizophrenia and autism. The difficulties of the AD patients in perceiving emotion may be considered secondary to their associated cognitive impairments (Koff et al. 1999). However, using the Florida Affect Battery (Bowers et al. 1998), an investigation found a significant difference between the AD and the healthy control groups on both emotion and cognition tasks (Bucks & Radford 2004). Another study suggested that impairments in emotional prosody processing in AD individuals may be ascribed to a more general prosodic processing impairment at the sentence level (Taler et al. 2008). Other studies found that emotional prosody perception by the AD group was significantly inferior to that of the control group, and insertion of linguistic content enlarged the difference (Horley et al. 2010).

Only a handful of studies on emotional prosody recognition in bipolar and depressive disorders are available; however, a meta-analysis of 16 studies from 2004 to 2019 also showed that these patients demonstrated overall deficits at a medium to large magnitude (see Tang et al. 2022 for a review). It was reported that the behavioral impairments in the processing of emotional prosody for patients with bipolar disorders could be related to early sensory impairments (see Van Rheenen & Rossell 2013 for a review), and in the case of depressed patients, the issues with emotional prosody were related to their cognitive (Uekermann et al. 2008) or sensory-perceptual (Emerson et al. 1999) deficiencies.

2.1.2. Recognition of linguistic prosody. Relative to the large number of studies of emotional recognition in mental disorders, fewer studies have investigated linguistic prosody. In this regard, ASD is an exception among the mental disorders due to the critical role of speech prosody in early language development (McCann et al. 2007). Individuals with ASD were reported to have difficulty perceiving the grammatical functions of lexical stress (Paul et al. 2005), syntactic chunking (Jarvinen-Pasley et al. 2008), question and statement intonation (Jarvinen-Pasley et al. 2008), and the pragmatic function of contrastive stress (Paul et al. 2005). In contrast, other studies showed that ASD participants were not inferior to their typical peers on tests of these same skills in lexical stress (Grossman et al. 2010), syntactic chunking (Paul et al. 2005), question and statement intonation (Paul et al. 2005), and contrastive stress (Globerson et al. 2015). Moreover, similar discrepancies across studies were also found on the pragmatic use of prosody, even though it is regarded as a hallmark weakness of ASD (Chevallier et al. 2011). One explanation is that pragmatic language abilities are selectively impaired in ASD. In addition, it is suggested that pragmatic abilities relying on assumptions about others' mental status (e.g., comprehension of irony, which requires inferring the intentions, beliefs, and feelings of others) rather than abilities limited to egocentric processing of context (e.g., interpretation of some indirect requests, which does not necessarily require assessment of the speaker's communicative intentions) were shown to be damaged (Deliens et al. 2018). This finding also implies that great efforts should be directed to developing appropriate prosodic tasks that can uncover the nature of communicative pragmatic difficulties in the ASD population.

Several studies investigated recognition of linguistic prosody in schizophrenia. However, no significant difference between patient and control groups was reported in a review of 11 studies between 1990 and 2019 (Lucarini et al. 2020). A detailed analysis of these studies revealed that the capacity to recognize grammatical prosody in patients with schizophrenia, such as resolving syntactic ambiguities or identifying question and statement intonations, was intact. However, specific impairments in identifying some pragmatic functions of prosody, such as detecting sarcasm and commands, were found in these patients. Lucarini et al. (2020) argued that the thin evidence for impairments in nonemotional prosody might be attributed to the high simplicity of the tasks, and they called for more attention to be directed toward pragmatic prosody investigation in this field.

Similarly, a review on AD patients reported conflicting results on linguistic prosody comprehension and found that most articles were of weak methodological quality (see Oh et al. 2021 for a review). Nonetheless, patients with major depressive disorders were found to be impaired in the comprehension of pragmatic prosody on emphatic stresses (Zurlo & Ruggiero 2021).

2.1.3. Summary of behavioral performance. With a brief review of the behavioral performances of individuals with mental disorders, we can conclude that recognition of emotional and linguistic prosody has been examined in all major mental illnesses. However, due to terminology confusion, many complex emotions, such as sarcasm and irony, have been referred to as emotional or affective prosody in some studies (see Icht et al. 2021 for a review), but as higher-order language functions (or pragmatic linguistic prosody) in others (see Mitchell & Crow 2005 for a review). Nevertheless, such complex emotions emerge as prosodic stimuli that can reveal significant differences in emotional and linguistic prosody recognition between patients and controls. Due to various experiment-related and participant-related factors, discrepant results exist. But the existing findings tend to show that prosodic tasks that represent complex emotions or pragmatic-social contexts can best identify impaired receptive prosody in individuals with mental disorders (Covington et al. 2005, Icht et al. 2021, Zurlo & Ruggiero 2021).

Except for general trends of impaired performances in receptive prosody, studies on this topic in mental disorders share several characteristics, which are crucial for our understanding. First, many studies examined the correlation of performances with higher-level cognitive abilities or with lower-level auditory skills. For example, the deficits of receptive prosody in individuals with mental disorders were reported to be significantly correlated with their impaired preattentive and attentive cognitive processing (Uekermann et al. 2008) or with abnormal sensory processing of pitch, intensity, or duration (Jahshan et al. 2013). Second, in combination with auditory behavioral tasks, different forms of stimuli (e.g., visual and auditory information) were involved in testing multisensory integration (Jones et al. 2011, Thaler et al. 2013). Third, neuroimaging methodology was used in uncovering the underlying mechanisms (Leitman et al. 2007, Wang et al. 2006). Finally, several investigations endeavored to use prosody recognition to identify individuals with mental problems before clinical symptoms were observed (Pawełczyk et al. 2018). For example, it was reported that ASD groups did not show the same preference for infant-directed speech as the typically developing (TD) group, which was related to a lack of neural enhancement effect in response to vowel exaggeration since TD children showed increased P1 amplitude toward hyperarticulated vowels while ASD children showed no sensitivity to this preattentive component of P1 (Chen et al. 2021). This phenomenon may indicate an early marker of ASD (see Filipe et al. 2018 for a review).

2.2. Theoretical Explanations

Receptive prosody research in mental disorders can provide implications for theories and models for cognitive and sensory processing of speech and emotion. At the same time, updated theoretical frameworks provide guidelines for further research work. Brain lateralization and the role of language in the processing of speech prosody are essential topics of concern among many central issues of theoretical explanations for behavioral performances of individuals with mental disorders.

2.2.1. Dynamic dual-stream model for prosody perception. Neurological interest in prosody was stimulated by the discovery that differently sited lesions have different effects on prosody. In early studies, behavioral experiments with brain-damaged subjects suggested that lexical stress prosody processing is a left hemisphere function, while the right hemisphere specializes in processing emotional prosody (Murphy & Cutting 1990). Later experiments further showed that the right hemisphere also mediates some language functions conveyed by pragmatic prosody. Impairments in these higher-order language functions displayed by patients with schizophrenia may significantly contribute to their social interaction deficits (see Mitchell & Crow 2005 for a review). The emotion-type hypothesis also posits that primary (basic) emotions are modulated by the right hemisphere, while social emotions are modulated by the left hemisphere (Ross 2021). However, neuroimaging experiments in healthy populations suggested that emotional prosody comprehension is subserved by bilateral mechanisms (Kotz et al. 2003, Schirmer & Kotz 2006). The conflicting results warrant further exploration with more neurophysiological and neuroimaging techniques in brain-damaged individuals (Leitman et al. 2007), which is also the trend in this research field (see Lin et al. 2018, Zhang et al. 2021 for reviews).

Despite the discrepancies, findings based on traditional lesion-based methods and modern functional imaging approaches on prosody perception have lent support to the dynamic dual-pathway model of auditory language comprehension. According to the model, segmental, lexical, and syntactic information is processed in the left hemisphere, while sentence-level suprasegmental information is processed in the right hemisphere (Friederici & Alter 2004, Frühholz et al. 2015, Sammler et al. 2015). This dynamic dual-stream model was then extended to show that, unlike nontonal language speakers, Mandarin Chinese speakers process lexical tones and sentence intonation primarily in the left hemisphere (Chien et al. 2021, Gandour et al. 2004). Cross-linguistic experiments further proved that speech and prosody perception relies upon a shared cortical auditory feature processing mechanism, which may be tuned to track and map the statistics (such as variance in pitch height, pitch contour, and speech rhythm) of a given language (Li et al. 2021). Emotional and linguistic prosody functions vary across listeners with different language and cultural backgrounds (Banse & Scherer 1996). In addition, as emotional prosody processing involves multistage brain networks (Grandjean 2021), the interplay of emotional and linguistic prosody functions at each stage in the pathway remains to be elucidated with more neuropsychological and cross-linguistic findings in the future.

2.2.2. Beyond the cognition-emotion-language trichotomy. Many behavioral and brain studies have been conducted to identify possible reasons for the impaired speech prosody comprehension in individuals with mental disorders. Higher-level dysfunctions in the network of theory of mind (ToM) (Baron-Cohen 1995) are one of the most explanatory theories that can be found in many studies of ASD and schizophrenia (Chevallier et al. 2011, Leitman et al. 2006, Razafimandimby et al. 2016, Rutherford et al. 2002). Cognitive deficit is another possible factor (Bucks & Radford 2004, Uekermann et al. 2008). But studies using higher-order language processes for accounting for impairments are scarce (Pawełczyk et al. 2018). Other investigations found an association of receptive prosody impairment with lower-level auditory dysfunction (Jahshan et al. 2013, Leitman et al. 2006). However, it was suggested that communication difficulties in autistic individuals are more consistent with the reduced social interest than the auditory dysfunction explanation (see Key & Slaboch 2021 for a review).

A meta-analysis of neuroimaging data across different social-cognitive tasks revealed that the higher level of neurocognitive processes involves the cognitive process, the affective process, and a combination of cognitive and affective processes and systems (see Schurz et al. 2021 for a review). Language, often regarded as a nonsocial term in linguistic research, was found to be embedded in the organizational properties of the motor system. However, as mentally disordered individuals suffer from a failure to use language in social interaction, we argue that integrating language into the interplay of emotion and cognition would facilitate a better understanding of their problems. Unlike the case for emotional prosody, there is no one-to-one relationship between pragmatic or attitudinal meanings of the utterance and their prosody. Interpreting such intentions depends on integrating multiple information from prosodic cues, linguistic context, and prior knowledge (Wichmann 2002). Attitudinal (pragmatic) prosody deserves equal attention as the widely studied emotional prosody (Mitchell & Ross 2013), even though developing pragmatic prosody tasks can be challenging. In addition, language serves as a context for emotion perception (Barrett et al. 2007) and plays an active role in constructing emotion concepts and shaping experience (Hoemann & Barrett 2019). For instance, the presence of emotional words in experimental tasks could have a significant influence over the neural representation of emotional experiences and perceptions (Brooks et al. 2017). All these findings have reminded us of a cautionary methodological note that an increase in linguistic complexity, such as the inclusion of emotion words in experimental tasks, could not be explained as simply an increase in the cognitive load, as was done in many studies. In this regard, understanding the dynamic neural interactions of emotion, cognition, and language is essential for selecting potent stimuli and developing effective training and treatment approaches to mitigate emotional dysfunction across various neuropsychiatric states (Liebenthal et al. 2016). Therefore, we propose going beyond the conventional trichotomy of cognition, emotion, and language for studying higher-level processing of speech prosody in basic and translational work; this approach can offer a practical framework for developing more appropriate methodologies in prosody research.

While it may not be difficult to manipulate certain acoustic features to represent target prosodic variations in research protocols suitable for the different populations of interest, considerations of linguistic influence should be taken into account in examining lower-level auditory processing. For example, mounting evidence has suggested that auditory processing impairments in ASD are more severe for speech than for nonspeech stimuli (see O'Connor 2012 for a review). Moreover, ASD children had reduced discrimination sensitivity for lexical tone contrast relative to TD peers, despite their equivalent or increased sensitivity to pitch contrasts of pure tones and hummed sounds (Yu et al. 2015, 2022). These findings also suggest that prosody serves as a bridge between lower-level acoustic elements and higher-level language components to help pinpoint the sources of problems to guide theoretical development and intervention practice.

2.3. Implications

In this section, after discussing the progress achieved on prosody recognition in mental disorders in recent years, we outline some of the critical implications for future research.

First, we propose a unified trinity of emotion-cognition-language for studying speech prosody recognition in neurotypicals and individuals with mental disorders. Complex emotions that can distinguish mentally disordered patients from their neurotypical controls demand the combined ability to process emotion, cognition, and language. Linguists need to work with experts in other related disciplines in developing more advanced methodologies with more comprehensive models and theories to examine the interactions across domains. One current trend is to examine multichannel integration, as prosody is never confined to the auditory domain. Prosody comprehension is usually accompanied by verbal semantic content, which results in multichannel processing. It is well acknowledged that both speech perception (Keough et al. 2019) and prosody perception (Pell et al. 2011) are multisensory processes. Moreover, social communication difficulties in mental disorders may involve deficits in cross-modal coordination (see Lin et al. 2020 and Zhang et al. 2022 for reviews). As reviewed above, individuals with mental disorders may hardly be differentiated from normal subjects by employing simple grammatical prosody or basic emotional prosody tasks. One plausible explanation is that the "disordered" individuals may have the intact ability in single-channel processing of basic prosodic category information but impaired skills in integrating processes across communication channels and sensory modalities. Therefore, multimodal assessment is also more practical for detecting impairments of pragmatic communication in individuals with mental disorders (Parola et al. 2021).

Second, to improve generalizability and explanatory power, the dynamic bilateral prosodic paths in the dual-stream model need to be further elaborated and updated to account for different functions of prosody, including the special characteristics of tonal languages. Further neuroimaging research on patients with mental disorders can help us understand the mechanisms of prosody perception at different levels and how different levels of processing may break down in various pathological or experimental conditions.

Third, current prosody evaluation tools are written mainly in English and several European languages. Given that pragmatic or affective prosody is language and culture dependent (Jiang et al. 2020), more tools need to be developed and customized to address linguistic/cultural diversity and equity. The same applies to scientific research. For instance, more investigations of tonal languages and Asian cultures can be helpful to obtain a comprehensive view of receptive prosody processing in individuals with mental disorders.

Finally, intervention and training programs that aim at pragmatic prosody recognition with materials and conditions tailored toward individual emotion, cognition, and language abilities are expected to be developed for ASD (Schreibman et al. 2015), schizophrenia (Hoekert et al. 2007), and other mental disorders. Although impairments in pragmatic or affective prosody recognition are systematically documented in research studies on individuals with various mental disorders (Icht et al. 2021, Lin et al. 2018, Oh et al. 2021), prosody comprehension/production tasks are rarely included in clinical settings.

3. EXPRESSION OF PROSODY IN MENTAL DISORDERS

Parallel to perceptual problems, the production of proper speech prosody is often compromised in individuals with mental disorders. The most prominent speech characteristics affected by early mental disorders are those related to abnormal prosody, which includes unusual fluctuations in time, pitch, loudness, and voice quality. These characteristic alterations underlying various mental states and conditions may be detected by analyzing the acoustic temporal and spectral parameters.

3.1. Behavioral Performances

In speech expression, people modulate their voices to accommodate various social contexts and motivational states so that linguistic prosody is intertwined with emotional prosody (Pisanski et al. 2016). However, how the linguistic and emotional production systems interact with each other remains poorly understood (Bryant 2021). In early studies, performances of emotional or linguistic prosody production were usually assessed by human raters (Cancelliere & Kertesz 1990, Murphy & Cutting 1990). However, with the increased capability and popularity of phonetic software, acoustic-phonetic analysis has been applied with automated approaches and more objective statistics (Alpert et al. 2001, Filipe et al. 2014). For instance, the incorporation of advanced machine learning algorithms makes it possible to implement automatic clustering/classification to assess prosodic features for detecting mental disorders in spontaneous speech (Haider et al. 2020, Yuan et al. 2021). In the following sections, we discuss the studies on expressive prosody using various approaches.

3.1.1. Subjective rating of expressive prosody. In many studies, the primary tasks in prosody expression were to ask the participants to imitate or read the sentences in one of the designated emotions or stress patterns, and trained judges were invited to rate the accuracy and intensity of the emotional expressions (Putnam & Kring 2007) or the accuracy of the lexical stress (Murphy & Cutting 1990). A meta-analysis on seven studies between 1980 and 2007 reported that the expressive ability of emotional prosody was impaired in schizophrenia with a large effect size (see Hoekert et al. 2007 for a review). But the findings were inconclusive due to the limited number of studies. Other studies reported that, even if individuals with mental disorders retained the ability to differentiate between emotions in production, they employed significantly different prosodic patterns than the controls (Horley et al. 2010, Hubbard et al. 2017).

Studies on the production of linguistic prosody in mental disorders have been scarce. Investigations have largely focused on acoustic parameters rather than ratings (Lucarini et al. 2020). For example, lexical analyses showed that changes in inflection and speech rate were associated with clinical ratings of negative symptoms of schizophrenia (Alpert et al. 2002, Cohen et al. 2008). There are some studies on expressive linguistic prosody with subjective ratings on several prosodic features in the ASD population. Due to high variability in methodology and participants, no consistent results on ASD individuals were observed in a review on 16 studies between 1980 and 2002 (see McCann & Peppe 2003 for a review). Some studies found significant differences between ASD and TD groups in producing grammatical stresses (McCann & Peppe 2003, Paul et al. 2005). Others reported that ASD individuals could demonstrate appropriately differentiated lexical stress patterns but showed atypically long productions (Grossman et al. 2010). In addition, children with ASD made significantly fewer nongrammatical pauses, suggesting that the stories told by autistic children reflected reduced cognitive and communicative demands (Thurber & Tager-Flusberg 1993). Although ASD individuals were reported to have difficulties in expressive linguistic prosody, pragmatic functions carrying social information and speakers' intentions were reported to be more impaired than were grammatical functions (Fine et al. 1991, Thurber & Tager-Flusberg 1993). However, variability in scoring criteria across studies and unquantified prosodic features in production may undermine the generalizability of the results (see McCann & Peppe 2003 for a review).

3.1.2. Phonetic acoustic analysis of expressive prosody. The advancement of analytic tools in speech acoustics has provided much objectivity to quantify the acoustic parameters of prosodic features (Banse & Scherer 1996). Even when individuals with mental disorders achieved usual accuracy with prosodic functions, acoustic analysis often revealed differences/alterations in prosodic form relative to matched peers (Loveall et al. 2021). The target acoustic parameters include pitch, duration, intensity, and voice quality. It was reported that the most common characteristics of abnormal speech for schizophrenia were reduced variations in intonation (monotone speech) as well as increased pauses, hesitations, and strained voice (Covington et al. 2005). AD patients showed reduced pitch variation and a slower speech rate relative to age-matched controls in producing emotional prosody (see Horley et al. 2010 for a review). Relative to the TD control group, ASD group members showed an exaggerated pitch range, atypical pitch contours, higher pitch variability (Filipe et al. 2014, Nadig & Shaw 2012, Sharda et al. 2010), and longer production (Diehl & Paul 2013, Filipe et al. 2014). Despite the findings on distinctive acoustic patterns associated with ASD, no single feature could serve as a marker for ASD or as a predictor of the severity of clinical features. While multivariate machine learning may be able to provide promising results about a cluster of biomarker features for ASD, more systematic cross-study validations are required (see Fusaroli et al. 2017 for a review).

3.1.3. Automatic classification of prosodic events for detection. Prosody production has been shown to signal an individual's specific cognitive and social functioning profile (Fusaroli et al. 2017). Changes in acoustic measures—especially duration, pitch, loudness, and voice quality that reflect the mental state—are well documented and substantially related to clinical observations and reports for mental health status examination (Alpert et al. 2001, Hirschberg et al. 2010). From a computational perspective, current machine learning methods can automatically extract purely acoustic features from spontaneous speech to identify mental illnesses with a fairly high detection success rate, like AD (Haider et al. 2020). Therefore, many machine learning studies have employed speech data to detect mental disorders such as depression, schizophrenia, and bipolar disorder (see Low et al. 2020 for a review). For example, prosodic and speech characteristics can be helpful in assessing depression and suicide risks (see Cummins et al. 2015 for a review). Vocal prosody could be a powerful tool to screen and monitor the progression and recovery of depressive disorders (Alpert et al. 2001, Yang et al. 2013). In addition, speech pause distribution and other acoustic parameters could be used as an early marker for AD (Meilan et al. 2014, Pastoriza-Domínguez et al. 2022, Yuan et al. 2021).

Along with acoustic parameters, some verbal or nonverbal vocalizations in speech also serve as pragmatic signals for social interactions (Bryant 2021). Filler disfluencies (e.g., "uh" and "um") may also serve as distinct discourse functions (Ward 2006), which could be employed to detect mental disorders such as ASD (Irvine et al. 2016, McGregor & Hadden 2020) and AD (Yuan et al. 2021).

3.1.4. Summary of behavioral performance.

In sum, expressive prosody in mental disorders shares some common characteristics. First, like receptive prosody, pragmatic prosody that reflects social communication appears to be more effective than linguistic prosody in distinguishing individuals with mental illnesses from their neurotypical controls. Second, data-driven automatic classification based on prosodic analysis can use sizeable spontaneous speech data for detecting and validating mental disorders but offers scarce information for our understanding of the role of prosody in speech communication and underlying causes. Third, as prosodic expression is multichannel and cross-modal, modeling acoustic details should be combined with semantic-syntactic components (Hirschberg et al. 2010, Tang et al. 2021) for speech prosody to be a better biomarker. If combined with other modalities, speech prosody can become a more valuable and beneficial diagnostic tool (Alghowinem et al. 2018). Finally, while machine learning is sensitive to detecting subtle acoustic signs indicative of neurocognitive disturbances, human evaluators in clinical diagnosis may not be able to perceive and utilize such nuances. The advantage of corpus-based automatic detection in not requiring the use of subjective rating scales or highly trained staff can be employed to help in preclinical screening of mental disorders. It is noteworthy that, in addition to problems with expressive prosody, patients and their unaffected first-degree relatives in ASD (Patel et al. 2019) and schizophrenia (Pawełczyk et al. 2018) showed higher-order language dysfunctions. These findings indicated that prosody comprehension and production tests could be instrumental to behavioral genetics in identifying individuals with a high risk of mental disorders such as ASD and schizophrenia.

3.2. Theoretical Explanations

Speech production is a complex neurocognitive and physiological process. It begins with cognitive and linguistic planning by forming the intended messages and retrieving corresponding words and phrases in the mental lexicon. This information is then stored briefly in working memory, where complex cognitive and linguistic functions and mapping operations at various levels of phonology, morphology, and syntax are performed and phonetic and prosodic representations are produced. In the final stage of motor action, a series of neuromuscular commands are executed to initiate the motoric actions for the sequenced sounds and words needed to utter the intended speech (Cummins et al. 2015). In the physiological part, speech production represents a complex motor behavior that requires accurate planning, sequencing, and coordination of the articulators forming the human vocal apparatus, which is accompanied by proprioceptive feedback and auditory feedback for output monitoring (Chrabaszcz et al. 2019, Kroger et al. 2009).

Impairments in speech production of individuals with mental disorders are explained in terms of higher-level neurocognitive dysfunctions as well as lower-level disruptions in motor-speech control. ToM is usually invoked to explain the impaired prosodic and other information processing of ASD and schizophrenic disorders to form inappropriate messages. For example, the unusual and improper use of intonation could be attributed to individuals not being aware of the communication value of specific intonation patterns. This deficiency may be explained by ToM in high-functioning autism (Fine et al. 1991, McCann et al. 2007). Moreover, a reduction in cognitive ability and working memory also affects speech planning, impairing neuromuscular motor coordination processes (Cummins et al. 2015). Such cognitive impairments and related abnormal speech traits have been found to be reliably detectable at mild stages of AD (Ivanova et al. 2022). There are significant correlations between the severity of mental disorders and pause-related measures in free speech production, showing that mentally disordered individuals have difficulty in discourse planning and choosing words (Alpert et al. 2001, Mundt et al. 2012). Motor-speech control is also affected by abnormal changes in affective states in mentally disordered patients (Cummins et al.

2015, Franich et al. 2021). Unusual mental conditions can disturb somatic and autonomic nervous systems, further influencing muscle tension and respiratory rates. Ultimately, these disturbances are expressed in the prosody and quality of the speech produced. Studies have shown that both prosodic and vocal source features are affected by a speaker's severity level of mental disorders such as depression (Moore et al. 2008). In addition, motor-speech control requires precise timing. In this regard, impairment of speech prosody in ASD may be linked with deficits in temporal processing (Franich et al. 2021).

To better understand the nature of mental health-related prosodic, articulatory, and acoustic alterations, future research needs to go beyond subjective ratings, acoustic feature analysis, and machine learning classification by tapping into mechanistic accounts with reference to neuroanatomical and physiological models of speech production. For example, the directions into velocities of articulators (DIVA) model is a well-established neural network model of speech motor skill acquisition and speech production. It describes how intended speech sounds are converted into articulatory movements, resulting in speech production (Golfinopoulos et al. 2010). As speech prosody is a property inherent in speech production, further research efforts could address how to accommodate the output of pitch-based prosodic cues into such a mechanism (Belyk & Brown 2014). In addition, the latest version of gradient order DIVA (GODIVA) model provides a mechanistic account of speech apraxia, a disorder of speech motor programming that can be found in pathological conditions such as ASD (Miller & Guenther 2021). It would be most fitting to integrate suprasegmental arrangement and modulation into the speech production model in the future; such a model could provide a testbed for behavioral and neuroimaging hypotheses about the underlying neural mechanisms responsible for motor programming in speakers with or without related pathological conditions.

3.3. Implications

In natural speech communication, linguistic prosody operates with emotional prosody in tandem to reveal one's mental state (Bryant 2021, Pisanski et al. 2016). Refined acoustic analyses of vocal modulation can be an excellent tool to distinguish patients with mental disorders from control participants (Fusaroli et al. 2017, Hirschberg et al. 2010). Many recent studies have demonstrated that machine learning algorithms based on acoustic feature analysis provide an objective, inexpensive, and noninvasive tool for detecting major mental disorders (Low et al. 2020, Pastoriza-Domínguez et al. 2022). However, several improvements are necessary to advance prosody investigation in this aspect.

First, challenges for corpus phonetics, including restricted availability of effective techniques and a lack of unified criteria (Liberman 2019), also apply to the construction and application of speech databases in mental disorders. Open-speech databases that researchers across disciplines can share are undoubtedly beneficial for all stakeholders. Efforts are needed to build large, publicly available speech corpora for mental disorders, following exemplar models such as TalkBank (MacWhinney 2007) and DementiaBank (Becker et al. 1994). In this regard, experts from multidisciplinary fields have recently formed a diverse international consortium of research in thought/language and communication (DISCOURSE) in psychosis to promote international collaboration efforts toward creating an open-source speech databank for psychosis (**https://discourseinpsychosis.org/**). However, sharing clinical data that may reveal personal identity or characteristics presents an ethical challenge (Insel 2018) that should be handled with caution.

Second, speech, as a measurable human behavior, provides a convenient and rich resource for identifying prosodic characteristics that can function as a biomarker for various mental disorders to facilitate early diagnosis, monitor disease progression, and quantify treatment progress. However, rigorous evaluations and standards are required for effective and safe use (Robin et al. 2020).

Third, current diagnosis of mental disorders primarily makes use of the acoustic analysis or subjective rating of prosodic and speech features (see Cummins et al. 2015, Low et al. 2020 for reviews). As mental disorder deficits are manifested across modalities and involve cross-modal coordination problems, one trend is to integrate speech prosodic features with verbal and visual information to generate better prediction results with multimodal databases (Alghowinem et al. 2018, Hirschberg et al. 2010, Tang et al. 2021).

Finally, individuals with mental disorders such as ASD and schizophrenia often display language impairments at phonological, morphological, syntactic, semantic, and pragmatic levels (Covington et al. 2005, Martin & McDonald 2003). Future work needs to incorporate systematic linguistic analysis with a focus on profiling language function, especially pragmatic language, into the explanation of expressive prosody impairments of patients with mental disorders (Martin & McDonald 2003, Zurlo & Ruggiero 2021). More theory-driven and model-based research is necessary to advance our understanding of the underlying mechanisms of abnormal speech prosody expression or vocal production in mental disorders (Fusaroli et al. 2017).

4. CONCLUSION

Speech prosody in mental disorders has increasingly received attention in various disciplines, including linguistics, psychology, psychiatry, neurology, and computer science. As a result, there are vast venues of research too broad to be covered in this review. As individuals with mental disorders fall into different categories and subcategories based on etiology and symptomatology, in-depth discussions with nuanced details for each specific disorder are beyond the scope of our review. Although we discuss prosody recognition and production independently here, more attention needs to be directed towards their binding relationship as shared neural systems to account for individual differences in socio-communicative ability (Aziz-Zadeh et al. 2010).

Our review of the recent developments in this research area has revealed three major themes. First, the prosodic deficits so frequently attributed to individuals with mental illnesses reside primarily in pragmatic and affective aspects of troubles in comprehending more complex emotions and communicative intentions, which requires an integrative ability to process emotion in the social interaction context. In this regard, an integrated theoretical framework of language, cognition, and emotion with a focus on cross-domain interactions is highly in need to explain the socio-communicative problems reflected in speech prosody comprehension and production. Without theoretical advancement and guidance, it is difficult to determine the relative independence as well as interconnections between language-related factors and cognitive and emotional factors to better characterize the high heterogeneity in individuals with psychiatric disorders in speech communication. As social emotions are generally language- and culture-dependent, studies on speech prosody in mental illnesses need to be more inclusive to cover both nontonal and tonal languages and culture-sensitive materials and settings. Second, prosody is a naturally multisensory and multimodal human behavior. With the advent of computational modeling and machine learning algorithms, proper construction of multimodal databases for normal and pathological samples will complement speech-based prosodic analysis and lead to more reliable quantitative biomarkers for early screening and the development of more effective intervention. Finally, multidisciplinary efforts are needed to move forward new frontiers of research and practice, with speech prosody regarded as a signal with tremendous social and linguistic significance and as a window into the speaker's mental state from a medical standpoint.

Successful collaborations across disciplines will inform and inspire new theoretically guided basic research and evidence-based clinical applications.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The writing of this review was supported by the Major Programs of the National Social Science Foundation of China (18ZDA293).

LITERATURE CITED

Alghowinem S, Goecke R, Wagner M, Epps J, Hyett M, et al. 2018. Multimodal depression detection: fusion analysis of paralinguistic, head pose and eye gaze behaviors. *IEEE Trans. Affect. Comput.* 9:478–90

Alpert M, Pouget ER, Silva RR. 2001. Reflections of depression in acoustic measures of the patient's speech. *J. Affect. Disord.* 66:59–69

Alpert M, Shaw RJ, Pouget ER, Lim KO. 2002. A comparison of clinical ratings with vocal acoustic measures of flat affect and alogia. *J. Psychiatr. Res.* 36:347–53

American Psychiatric Association. 1994. *Diagnostic and Statistical Manual of Mental Disorders, 4th Ed*. Washington, DC: Am. Psychiatr. Publ.

Angeleri R, Bosco FM, Zettin M, Sacco K, Colle L, Bara BG. 2008. Communicative impairment in traumatic brain injury: a complete pragmatic assessment. *Brain Lang.* 107:229–45

Aziz-Zadeh L, Sheng T, Gheytanchi A. 2010. Common premotor regions for the perception and production of prosody and correlations with empathy and prosodic ability. *PLOS ONE* 5:e8759

Bambini V, Arcara G, Bechi M, Buonocore M, Cavallaro R, Bosia M. 2016. The communicative impairment as a core feature of schizophrenia: frequency of pragmatic deficit, cognitive substrates, and relation with quality of life. *Compr. Psychiatry* 71:106–20

Banse R, Scherer KR. 1996. Acoustic profiles in vocal emotion expression. *J. Personal. Soc. Psychol.* 70:614–36

Baron-Cohen S. 1995. *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press

Barrett LF, Lindquist KA, Gendron M. 2007. Language as context for the perception of emotion. *Trends Cogn. Sci.* 11:327–32

Baxter AJ, Brugha TS, Erskine HE, Scheurer RW, Vos T, Scott JG. 2015. The epidemiology and global burden of autism spectrum disorders. *Psychol. Med.* 45:601–13

Becker JT, Boller F, Lopez OL, Saxton J, McGonigle KL. 1994. The natural history of Alzheimer's disease. Description of study cohort and accuracy of diagnosis. *Arch. Neurol.* 51:585–94

Belyk M, Brown S. 2014. Perception of affective and linguistic prosody: an ALE meta-analysis of neuroimaging studies. *Soc. Cogn. Affect. Neurosci.* 9:1395–403

Benedetti V, Weill-Chounlamountry A, Pradat-Diehl P, Villain M. 2021. Assessment tools and rehabilitation treatments for aprosodia following acquired brain injury: a scoping review. *Int. J. Lang. Commun. Dis.* 57:474–96

Bolinger D. 1989. *Intonation and Its Uses: Melody in Grammar and Discourse*. Stanford, CA: Stanford Univ. Press

Bowers D, Blonder L, Heilman K. 1998. *Florida affect battery*. Manual, Cent. Neuropsychol. Stud., Dep. Neurol., Univ. Fla.

Brooks JA, Shablack H, Gendron M, Satpute AB, Parrish MH, Lindquist KA. 2017. The role of language in the experience and perception of emotion: a neuroimaging meta-analysis. *Soc. Cogn. Affect. Neurosci.* 12:169–83

Bryant GA. 2021. The evolution of human vocal emotion. *Emot. Rev.* 13:25–33

Bucks RS, Radford SA. 2004. Emotion processing in Alzheimer's disease. *Aging Ment. Health* 8:222–32

Cancelliere AE, Kertesz A. 1990. Lesion localization in acquired deficits of emotional expression and comprehension. *Brain Cogn.* 13:133–47

Carvalho AF, Solmi M, Sanches M, Machado MO, Stubbs B, et al. 2020. Evidence-based umbrella review of 162 peripheral biomarkers for major mental disorders. *Transl. Psychiatry* 10:152

Chen F, Zhang H, Ding HW, Wang SP, Peng G, Zhang Y. 2021. Neural coding of formant-exaggerated speech and nonspeech in children with and without autism spectrum disorders. *Autism Res*. 14:1357–74

Chevallier C, Noveck I, Happé F, Wilson D. 2011. What's in a voice? Prosody as a test case for the Theory of Mind account of autism. *Neuropsychologia* 49:507–17

Chien PJ, Friederici AD, Hartwigsen G, Sammler D. 2021. Intonation processing increases task-specific fronto-temporal connectivity in tonal language speakers. *Hum. Brain Mapp.* 42:161–74

Chrabaszcz A, Neumann W-J, Stretcu O, Lipski WJ, Bush A, et al. 2019. Subthalamic nucleus and sensorimotor cortex activity during speech production. *J. Neurosci.* 39:2698–708

Cohen AS, Alpert M, Nienow TM, Dinzeo TJ, Docherty NM. 2008. Computerized measurement of negative symptoms in schizophrenia. *J. Psychiatr. Res.* 42:827–36

Covington MA, He CZ, Brown C, Naci L, McClain JT, et al. 2005. Schizophrenia and the structure of language: the linguist's view. *Schizophr. Res.* 77:85–98

Crystal D. 1986. Prosodic development. In *Language Acquisition*, ed. P Fletcher, M Garman, pp. 33–48. Cambridge, UK: Cambridge Univ. Press

Cummins N, Scherer S, Krajewski J, Schnieder S, Epps J, Quatieri TF. 2015. A review of depression and suicide risk assessment using speech analysis. *Speech Commun*. 71:10–49

Cutler A, Dahan D, van Donselaar W. 1997. Prosody in the comprehension of spoken language: a literature review. *Lang. Speech* 40:141–201

Cutler A, Pearson M. 2018. On the analysis of prosodic turn-taking cues. In *Intonation in Discourse*, ed. C Johns-Lewis, pp. 139–55. Abingdon, UK: Routledge

Deliens G, Papastamou F, Ruytenbeek N, Geelhand P, Kissine M. 2018. Selective pragmatic impairment in autism spectrum disorder: indirect requests versus irony. *J. Autism Dev. Disord.* 48:2938–52

Diehl JJ, Paul R. 2013. Acoustic and perceptual measurements of prosody production on the profiling elements of prosodic systems in children by children with autism spectrum disorders. *Appl. Psycholinguist.* 34:135–61

Edwards J, Jackson HJ, Pattison PE. 2002. Emotion recognition via facial expression and affective prosody in schizophrenia: a methodological review. *Clin. Psychol. Rev.* 22:789–832

Ekman P. 1992. An argument for basic emotions. *Cogn. Emot.* 6:169–200

Emerson CS, Harrison DW, Everhart DE. 1999. Investigation of receptive affective prosodic ability in school-aged boys with and without depression. *Neuropsychiatry Neuropsychol. Behav. Neurol.* 12:102–9

Fairbanks G, Pronovost W. 1938. Vocal pitch during simulated emotion. *Science* 88:382–83

Filipe MG, Frota S, Castro SL, Vicente SG. 2014. Atypical prosody in Asperger syndrome: perceptual and acoustic measurements. *J. Autism Dev. Disord.* 44:1972–81

Filipe MG, Watson L, Vicente SG, Frota S. 2018. Atypical preference for infant-directed speech as an early marker of autism spectrum disorders? A literature review and directions for further research. *Clin. Linguist. Phon.* 32:213–31

Fine J, Bartolucci G, Ginsberg G, Szatmari P. 1991. The use of intonation to communicate in pervasive developmental disorders. *J. Child Psychol. Psychiatry* 32:771–82

Franich K, Wong HY, Yu ACL, To CKS. 2021. Temporal coordination and prosodic structure in autism spectrum disorder: timing across speech and non-speech motor domains. *J. Autism Dev. Disord.* 51:2929–49

Friederici AD, Alter K. 2004. Lateralization of auditory language functions: a dynamic dual pathway model. *Brain Lang*. 89:267–76

Froming K, Levy M, Schaffer S, Ekman P. 2022. *The Comprehensive Affect Testing System*. **https://www.millisecond.com/download/library/cats**

Frühholz S, Gschwind M, Grandjean D. 2015. Bilateral dorsal and ventral fiber pathways for the processing of affective prosody identified by probabilistic fiber tracking. *Neuroimage* 109:27–34

Fusaroli R, Lambrechts A, Bang D, Bowler DM, Gaigg SB. 2017. Is voice a marker for autism spectrum disorder? A systematic review and meta-analysis. *Autism Res*. 10:384–407

Gandour J, Tong YX, Wong D, Talavage T, Dzemidzic M, et al. 2004. Hemispheric roles in the perception of speech prosody. *Neuroimage* 23:344–57

Globerson E, Amir N, Kishon-Rabin L, Golan O. 2015. Prosody recognition in adults with high-functioning autism spectrum disorders: from psychoacoustics to cognition. *Autism Res*. 8:153–63

Golan O, Baron-Cohen S, Hill J. 2006. The Cambridge Mindreading (CAM) Face-Voice Battery: testing complex emotion recognition in adults with and without Asperger syndrome. *J. Autism Dev. Disord.* 36:169–83

Golfinopoulos E, Tourville JA, Guenther FH. 2010. The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *Neuroimage* 52:862–74

Grandjean D. 2021. Brain networks of emotional prosody processing. *Emot. Rev.* 13:34–43

Grossman RB, Bemis RH, Skwerer DP, Tager-Flusberg H. 2010. Lexical and affective prosody in children with high-functioning autism. *J. Speech Lang. Hear. Res.* 53:778–93

Haider F, de la Fuente S, Luz S. 2020. An assessment of paralinguistic acoustic features for detection of Alzheimer's dementia in spontaneous speech. *IEEE J. Sel. Top. Signal Process.* 14:272–81

Hampel H, Hardy J, Blennow K, Chen C, Perry G, et al. 2021. The amyloid-beta pathway in Alzheimer's disease. *Mol. Psychiatry* 26:5481–503

Hirschberg J, Hjalmarsson A, Elhadad N. 2010. "You're as sick as you sound": using computational approaches for modeling speaker state to gauge illness and recovery. In *Advances in Speech Recognition*, ed. A Neustein, pp. 305–22. Boston: Springer

Hoekert M, Kahn RS, Pijnenborg M, Aleman A. 2007. Impaired recognition and expression of emotional prosody in schizophrenia: review and meta-analysis. *Schizophr. Res.* 96:135–45

Hoemann K, Barrett LF. 2019. Concepts dissolve artificial boundaries in the study of emotion and cognition, uniting body, brain, and mind. *Cogn. Emot.* 33:67–76

Horley K, Reid A, Burnham D. 2010. Emotional prosody perception and production in dementia of the Alzheimer's type. *J. Speech Lang. Hear. Res.* 53:1132–46

Hubbard DJ, Faso DJ, Assmann PF, Sasson NJ. 2017. Production and perception of emotional prosody by adults with autism spectrum disorder. *Autism Res*. 10:1991–2001

Icht M, Zukerman G, Ben-Itzchak E, Ben-David BM. 2021. Keep it simple: identification of basic versus complex emotions in spoken language in individuals with autism spectrum disorder without intellectual disability: a meta-analysis study. *Autism Res*. 14:1948–64

Insel TR. 2018. Digital phenotyping: a global tool for psychiatry. *World Psychiatry* 17:276–77

Irvine CA, Eigsti IM, Fein DA. 2016. Uh, um, and autism: filler disfluencies as pragmatic markers in adolescents with optimal outcomes from autism spectrum disorder. *J. Autism Dev. Disord.* 46:1061–70

Ivanova O, Meilán JJG, Martínez-Sánchez F, Martínez-Nicolás I, Llorente TE, González NC. 2022. Discriminating speech traits of Alzheimer's disease assessed through a corpus of reading task for Spanish language. *Comput. Speech Lang.* 73:101341

Jahshan C, Wynn JK, Green MF. 2013. Relationship between auditory processing and affective prosody in schizophrenia. *Schizophr. Res.* 143:348–53

Jarvinen-Pasley A, Peppe S, King-Smith G, Heaton P. 2008. The relationship between form and function level receptive prosodic abilities in autism. *J. Autism Dev. Disord.* 38:1328–40

Jiang X, Gossack-Keenan K, Pell MD. 2020. To believe or not to believe? How voice and accent information in speech alter listener impressions of trust. *Q. J. Exp. Psychol.* 73:55–79

Joanette Y, Ska B, Côté H. 2004. *Protocole Montréal d'évaluation de la communication*. Tool, IUGM, Montreal. **https://criugm.qc.ca/outils/protocole-montreal-devaluation-mec/**

Jones CRG, Pickles A, Falcaro M, Marsden AJS, Happe F, et al. 2011. A multimodal approach to emotion recognition ability in autism spectrum disorders. *J. Child Psychol. Psychiatry* 52:275–85

Keough M, Derrick D, Gick B. 2019. Cross-modal effects in speech perception. *Annu. Rev. Linguist.* 5:49–66

Key AP, Slaboch KD. 2021. Speech processing in autism spectrum disorder: an integrative review of auditory neurophysiology findings. *J. Speech Lang. Hear. Res.* 64:4192–212

Klein-Koerkamp Y, Beaudoin M, Baciu M, Hot P. 2012. Emotional decoding abilities in Alzheimer's disease: a meta-analysis. *J. Alzheimer's Dis.* 32:109–25

Koff E, Zaitchik D, Montepare J, Albert MS. 1999. Emotion processing in the visual and auditory domains by patients with Alzheimer's disease. *J. Int. Neuropsychol. Soc.* 5:32–40

Kotz SA, Meyer M, Alter K, Besson M, von Cramon DY, Friederici AD. 2003. On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain Lang*. 86:366–76

Kroger BJ, Kannampuzha J, Neuschaefer-Rube C. 2009. Towards a neurocomputational model of speech production and perception. *Speech Commun*. 51:793–809

Ladd DR. 2008. *Intonational Phonology*. Cambridge, UK: Cambridge Univ. Press

Ladd DR Jr. 1980. *The Structure of Intonational Meaning: Evidence from English*. Bloomington, IN: Indiana Univ. Press

Lartseva A, Dijkstra T, Buitelaar JK. 2015. Emotional language processing in autism spectrum disorders: a systematic review. *Front. Hum. Neurosci.* 8:991

Leitman DI, Hoptman MJ, Foxe JJ, Saccente E, Wylie GR, et al. 2007. The neural substrates of impaired prosodic detection in schizophrenia and its sensorial antecedents. *Am. J. Psychiatry* 164:474–82

Leitman DI, Ziwich R, Pasternak R, Javitt DC. 2006. Theory of Mind (ToM) and counterfactuality deficits in schizophrenia: misperception or misinterpretation? *Psychol. Med.* 36:1075–83

Li YN, Tang C, Lu JF, Wu JS, Chang EF. 2021. Human cortical encoding of pitch in tonal and non-tonal languages. *Nat. Commun.* 12:1161

Liberman MY. 2019. Corpus phonetics. *Annu. Rev. Linguist.* 5:91–107

Liebenthal E, Silbersweig DA, Stern E. 2016. The language, tone and prosody of emotions: neural substrates and dynamics of spoken-word emotion perception. *Front. Neurosci.* 10:506

Lin Y, Ding H, Zhang Y. 2018. Emotional prosody processing in schizophrenic patients: a selective review and meta-analysis. *J. Clin. Med.* 7:363

Lin Y, Ding H, Zhang Y. 2020. Multisensory integration of emotion in schizophrenic patients. *Multisens. Res.* 33:865–901

Loveall SJ, Hawthorne K, Gaines M. 2021. A meta-analysis of prosody in autism, Williams syndrome, and Down syndrome. *J. Commun. Disord.* 89:106055

Low DM, Bentley KH, Ghosh SS. 2020. Automated assessment of psychiatric disorders using speech: a systematic review. *Laryngoscope Investig. Otolaryngol.* 5:96–116

Lucarini V, Grice M, Cangemi F, Zimmermann JT, Marchesi C, et al. 2020. Speech prosody as a bridge between psychopathology and linguistics: the case of the schizophrenia spectrum. *Front. Psychiatry* 11:531863

MacWhinney B. 2007. The Talkbank Project. In *Creating and Digitizing Language Corpora*, Vol. 1: *Synchronic Databases*, ed. JC Beal, KP Corrigan, HL Moisl, pp. 163–80. London: Palgrave Macmillan

Martin I, McDonald S. 2003. Weak coherence, no theory of mind, or executive dysfunction? Solving the puzzle of pragmatic language disorders. *Brain Lang*. 85:451–66

McCann J, Peppe S. 2003. Prosody in autism spectrum disorders: a critical review. *Int. J. Lang. Commun. Dis.* 38:325–50

McCann J, Peppe S, Gibbon FE, O'Hare A, Rutherford M. 2007. Prosody and its relationship to language in school-aged children with high-functioning autism. *Int. J. Lang. Commun. Dis.* 42:682–702

McCutcheon RA, Reis Marques T, Howes OD. 2020. Schizophrenia: an overview. *JAMA Psychiatry* 77:201–10

McGregor KK, Hadden RR. 2020. Brief report: "Um" fillers distinguish children with and without ASD. *J. Autism Dev. Disord.* 50:1816–21

Meilan JJG, Martinez-Sanchez F, Carro J, Lopez DE, Millian-Morell L, Arana JM. 2014. Speech in Alzheimer's disease: Can temporal and acoustic parameters discriminate dementia? *Dement. Geriatr. Cogn. Disord.* 37:327–34

Miller HE, Guenther FH. 2021. Modelling speech motor programming and apraxia of speech in the DIVA/GODIVA neurocomputational framework. *Aphasiology* 35:424–41

Mitchell RLC, Crow TJ. 2005. Right hemisphere language functions and schizophrenia: the forgotten hemisphere? *Brain* 128:963–78

Mitchell RLC, Ross ED. 2013. Attitudinal prosody: what we know and directions for future study. *Neurosci. Biobehav. Rev.* 37:471–79

Monrad-Krohn GH. 1947. The prosodic quality of speech and its disorders. *Acta Psychiatr. Scand.* 22:255–69

Moore E II, Clements MA, Peifer JW, Weisser L. 2008. Critical analysis of the impact of glottal features in the classification of clinical depression in speech. *IEEE Trans. Biomed. Eng.* 55:96–107

Mundt JC, Vogel AP, Feltner DE, Lenderking WR. 2012. Vocal acoustic biomarkers of depression severity and treatment response. *Biol. Psychiatry* 72:580–87

Murphy D, Cutting J. 1990. Prosodic comprehension and expression in schizophrenia. *J. Neurol. Neurosurg. Psychiatry* 53:727–30

Nadig A, Shaw H. 2012. Acoustic and perceptual measurement of expressive prosody in high-functioning autism: increased pitch range and what it means to listeners. *J. Autism Dev. Disord.* 42:499–511

O'Connor K. 2012. Auditory processing in autism spectrum disorder: a review. *Neurosci. Biobehav. Rev.* 36:836–54

Oh C, Morris RJ, Wang X. 2021. A systematic review of expressive and receptive prosody in people with dementia. *J. Speech Lang. Hear. Res.* 64:3803–25

Parola A, Gabbatore I, Berardinelli L, Salvini R, Bosco FM. 2021. Multimodal assessment of communicative-pragmatic features in schizophrenia: a machine learning approach. *NPJ Schizophr.* 7:28

Pastoriza-Domínguez P, Torre IG, Diéguez-Vide F, Gómez-Ruiz I, Geladó S, et al. 2022. Speech pause distribution as an early marker for Alzheimer's disease. *Speech Commun*. 136:107–17

Patel SP, Kim JH, Larson CR, Losh M. 2019. Mechanisms of voice control related to prosody in autism spectrum disorder and first-degree relatives. *Autism Res*. 12:1192–210

Paul R, Augustyn A, Klin A, Volkmar FR. 2005. Perception and production of prosody by speakers with autism spectrum disorders. *J. Autism Dev. Disord.* 35:205–20

Pawełczyk A, Łojek E, Żurner N, Gawłowska-Sawosz M, Pawełczyk T. 2018. Higher-order language dysfunctions as a possible neurolinguistic endophenotype for schizophrenia: evidence from patients and their unaffected first degree relatives. *Psychiatry Res*. 267:63–72

Pell MD. 2006. Judging emotion and attitudes from prosody following brain damage. *Prog. Brain Res.* 156:303–17

Pell MD, Jaywant A, Monetta L, Kotz SA. 2011. Emotional speech processing: disentangling the effects of prosody and semantic cues. *Cogn. Emot.* 25:834–53

Pell MD, Paulmann S, Dara C, Alasseri A, Kotz SA. 2009. Factors in the recognition of vocally expressed emotions: a comparison of four languages. *J. Phon.* 37:417–35

Pierrehumbert JB, Hirschberg J. 1990. The meaning of intonational contours in the interpretation of discourse. In *Intentions in Communication*, ed. P Cohen, J Morgan, M Pollack, pp. 271–311. Cambridge, MA: MIT Press

Pisanski K, Cartei V, McGettigan C, Raine J, Reby D. 2016. Voice modulation: a window into the origins of human vocal control? *Trends Cogn. Sci.* 20:304–18

Putnam KM, Kring AM. 2007. Accuracy and intensity of posed emotional expressions in unmedicated schizophrenia patients: vocal and facial channels. *Psychiatry Res*. 151:67–76

Razafimandimby A, Herve PY, Marzloff V, Brazo P, Tzourio-Mazoyer N, Dollfus S. 2016. Functional deficit of the medial prefrontal cortex during emotional sentence attribution in schizophrenia. *Schizophr. Res.* 178:86–93

Robin J, Harrison JE, Kaufman LD, Rudzicz F, Simpson W, Yancheva M. 2020. Evaluation of speech-based digital biomarkers: review and recommendations. *Digit. Biomark*. 4:99–108

Ross ED. 2021. Differential hemispheric lateralization of emotions and related display behaviors: emotion-type hypothesis. *Brain Sci*. 11:1034

Ross ED, Thompson RD, Yenkosky J. 1997. Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions. *Brain Lang*. 56:27–54

Rutherford MD, Baron-Cohen S, Wheelwright S. 2002. Reading the mind in the voice: a study with normal adults and adults with Asperger syndrome and high functioning autism. *J. Autism Dev. Disord.* 32:189–94

Sammler D, Grosbras MH, Anwander A, Bestelmeyer PEG, Belin P. 2015. Dorsal and ventral pathways for prosody. *Curr. Biol.* 25:3079–85

Schirmer A, Kotz SA. 2006. Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn. Sci.* 10:24–30

Schreibman L, Dawson G, Stahmer AC, Landa R, Rogers SJ, et al. 2015. Naturalistic developmental behavioral interventions: empirically validated treatments for autism spectrum disorder. *J. Autism Dev. Disord.* 45:2411–28

Schurz M, Radua J, Tholen MG, Maliske L, Margulies DS, et al. 2021. Toward a hierarchical model of social cognition: a neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychol. Bull.* 147:293–327

Science. 2021. *125 questions: exploration and discovery*. Booklet, *Science*, Washington, DC
Sharda M, Subhadra TP, Sahay S, Nagaraja C, Singh L, et al. 2010. Sounds of melody-pitch patterns of speech in autism. *Neurosci. Lett.* 478:42–45
Shriberg LD, Paul R, McSweeny JL, Klin A, Cohen DJ, Volkmar FR. 2001. Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *J. Speech Lang. Hear. Res.* 44:1097–115
Taler V, Chertkow H, Baum SR, Saumier D. 2008. Comprehension of grammatical and emotional prosody is impaired in Alzheimer's disease. *Neuropsychology* 22:188–95
Tang E, Zhang M, Chen Y, Lin Y, Ding H. 2022. Recognition of affective prosody in bipolar and depressive conditions: a systematic review and meta-analysis. *J. Affect. Disord.* 313:126–36
Tang SX, Kriz R, Cho S, Park SJ, Harowitz J, et al. 2021. Natural language processing methods are sensitive to sub-clinical linguistic differences in schizophrenia spectrum disorders. *NPJ Schizophr*. 7:25
Thaler NS, Strauss GP, Sutton GP, Vertinski M, Ringdahl EN, et al. 2013. Emotion perception abnormalities across sensory modalities in bipolar disorder with psychotic features and schizophrenia. *Schizophr. Res.* 147:287–92
Thurber C, Tager-Flusberg H. 1993. Pauses in the narratives produced by autistic, mentally retarded, and normal children as an index of cognitive demand. *J. Autism Dev. Disord.* 23:309–22
Uekermann J, Abdel-Hamid M, Lehmkaemper C, Vollmoeller W, Daum I. 2008. Perception of affective prosody in major depression: a link to executive functions? *J. Int. Neuropsychol. Soc.* 14:552–61
Van Rheenen TE, Rossell SL. 2013. Is the non-verbal behavioural emotion-processing profile of bipolar disorder impaired? A critical review. *Acta Psychiatr. Scand.* 128:163–78
Wang AT, Lee SS, Sigman M, Dapretto M. 2006. Neural basis of irony comprehension in children with autism: the role of prosody and context. *Brain* 129:932–43
Ward N. 2006. Non-lexical conversational sounds in American English. *Pragmat. Cogn.* 14:129–82
Wichmann A. 2002. *Attitudinal intonation and the inferential process*. Presented at International Conference on Speech Prosody, Aix-en-Provence, Fr., April 11–13
World Health Organization. 2022. *Mental disorders*. Fact Sheet, World Health Organization. **https://www.who.int/news-room/fact-sheets/detail/mental-disorders**
Wu TC, Jia XQ, Shi HF, Niu JQ, Yin XH, et al. 2021. Prevalence of mental health problems during the COVID-19 pandemic: a systematic review and meta-analysis. *J. Affect. Disord.* 281:91–98
Yang Y, Fairbairn C, Cohn JF. 2013. Detecting depression severity from vocal prosody. *IEEE Trans. Affect. Comput.* 4:142–50
Yu LD, Fan YB, Deng ZZ, Huang D, Wang SP, Zhang Y. 2015. Pitch processing in tonal-language-speaking children with autism: an event-related potential study. *J. Autism Dev. Disord.* 45:3656–67
Yu LD, Huang D, Wang SP, Zhang Y. 2022. Reduced neural specialization for word-level linguistic prosody in children with autism. *J. Autism Dev. Disord.* **https://doi.org/10.1007/s10803-022-05720-x**
Yuan JH, Cai XY, Bian YC, Ye Z, Church K. 2021. Pauses for detection of Alzheimer's disease. *Front. Comput. Sci.* 2:624624488
Zhang M, Chen Y, Lin Y, Ding H, Zhang Y. 2022. Multi-channel perception of emotion in speech, voice, facial expression and gesture in individuals with autism: a scoping review. *J. Speech Lang. Hear. Res.* 65:1435–49
Zhang M, Xu S, Chen Y, Lin Y, Ding H, Zhang Y. 2021. Recognition of affective prosody in autism spectrum conditions: a systematic review and meta-analysis. *Autism* 26:798–813
Zurlo MC, Ruggiero M. 2021. Assessing pragmatic language skills in adults with major depressive disorder: an exploratory study. *Psychopathology* 54:78–91